DOI: 10.1534/genetics.104.032821

Evolution of Recombination Due to Random Drift

N. H. Barton* and Sarah P. Otto†,[1]

**School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom and* †*Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4*

Manuscript received June 24, 2004
Accepted for publication January 10, 2005

ABSTRACT

In finite populations subject to selection, genetic drift generates negative linkage disequilibrium, on average, even if selection acts independently (*i.e.*, multiplicatively) upon all loci. Negative disequilibrium reduces the variance in fitness and hence, by Fisher's (1930) fundamental theorem, slows the rate of increase in mean fitness. Modifiers that increase recombination eliminate the negative disequilibria that impede selection and consequently increase in frequency by "hitchhiking." Thus, stochastic fluctuations in linkage disequilibrium in finite populations favor the evolution of increased rates of recombination, even in the absence of epistatic interactions among loci and even when disequilibrium is initially absent. The method developed within this article allows us to quantify the strength of selection acting on a modifier allele that increases recombination in a finite population. The analysis indicates that stochastically generated linkage disequilibria do select for increased recombination, a result that is confirmed by Monte Carlo simulations. Selection for a modifier that increases recombination is highest when linkage among loci is tight, when beneficial alleles rise from low to high frequency, and when the population size is small.

RECOMBINATION breaks down the linkage disequilibrium generated between genes. In the absence of mutation and in a constant environment, the linkage disequilibrium present in a population will largely be a product of selection favoring specific gene combinations. In this case, one would expect that recombination would be detrimental and that genes that modify recombination rates would evolve to reduce recombination. This intuition is confirmed by the "reduction principle": in models at equilibrium under viability selection, modifier alleles that reduce recombination are favored in the absence of mutation (Feldman 1972; Feldman *et al.* 1980; Altenberg and Feldman 1987). How then is recombination maintained at appreciable levels in most higher organisms? In this article, we explore one possibility: that recombination is favored because it reduces the linkage disequilibrium generated stochastically in a population by random genetic drift.

For recombination to have an influence on the process of evolution, linkage disequilibria must exist within the genome. Although disequilibria can be generated by stochastic as well as deterministic forces (Felsenstein 1988; Kondrashov 1993), most work on the evolution of recombination has focused on deterministic models that assume an infinite population size. In a single unstructured population of infinite size, the spread of modifier alleles that increase recombination has been observed under three selective regimes. When epistatic interactions fluctuate over time, recombination can be favored, but only if fluctuations occur over a timescale of a few generations, such that linkage disequilibria that were initially advantageous soon become disadvantageous (Charlesworth 1976, 1990; Maynard Smith 1978; Barton 1995a). When there is directional selection, increased recombination rates can evolve under a wider range of conditions: essentially, there must be weak negative epistasis among favorable alleles (Maynard Smith 1980, 1988; Charlesworth 1993; Barton 1995a). Weak negative epistasis implies that the fitness of a genotype with several advantageous alleles is slightly less than would be expected from the product of the fitnesses of each allele measured separately. In this case, selection generates negative linkage disequilibrium between favorable alleles, which reduces the genetic variance present in the population. Recombination destroys this disequilibrium, increasing the variance in the population and improving the response of the population to selection. Finally, purifying selection against recurrent deleterious mutations also favors the evolution of recombination (Feldman *et al.* 1980; Kondrashov 1982, 1984, 1988; Barton 1995a; Otto and Feldman 1997). Again, increased recombination is favored if there are weak negative epistatic interactions among mutations, implying that the mutations interact synergistically to cause a further than expected reduction in fitness.

While negative epistasis can provide the raw material (negative linkage disequilibrium) for recombination to be favored at a modifier locus, it is not clear why epistasis should generally be negative—that is, why fitness interactions should impede selection. Experimental evi-

[1]*Corresponding author:* Department of Zoology, University of British Columbia, 6270 University Blvd., Vancouver, BC, Canada V6T 1Z4. E-mail: otto@zoology.ubc.ca

dence for negative epistasis among random mutations is limited (Rice 2002). Although negative epistasis has been observed in some cases (*e.g.*, Mukai 1969 in *Drosophila melanogaster*; de Visser *et al.* 1996 in *Chlamydomonas moewusii*; but see West *et al.* 1998), positive epistasis has been observed in other cases (*e.g.*, Seager and Ayala 1982, Seager *et al.* 1982 in *D. melanogaster*; but see Charlesworth 1990). Further studies in *Escherichia coli* (Elena and Lenski 1997) and *Aspergillus niger* (de Visser *et al.* 1997) have found significant evidence for both positive and negative epistasis but without a preponderance of either form. On theoretical grounds, negative epistasis has been invoked to account for the survival of populations in the face of a high rate of deleterious mutations. In the absence of epistasis, mean fitness would be $\exp(-U)$, where U is the genome-wide rate of deleterious mutations. Thus, mean fitness would be quite low in organisms with a high genomic mutation rate. The mean fitness of a population can be increased, however, in sexual populations if deleterious mutations interact synergistically (*i.e.*, exhibit negative epistasis; Kimura and Maruyama 1966; Kondrashov 1982, 1988). Yet synergistic interactions need not be invoked to explain the persistence of populations if the rate of deleterious mutation is low (Keightley 1997; Keightley and Caballero 1997; Keightley and Eyre-Walker 2000), if compensatory mutations are common, or if selection is "soft" such that the mean fitness of a population has little impact on the number of surviving individuals (Turner and Williamson 1968). Another approach has been to try to predict the form of epistasis, *a priori*, from models of the underlying biological processes affecting fitness. For example, using metabolic control theory, Szathmary (1993) found that negative epistasis among mutations predominates when flux through a metabolic pathway is under stabilizing selection but that positive epistasis predominates when maximal flux is favored. At this stage, both the experimental evidence and theoretical arguments concerning the nature of interactions among mutations remain equivocal and form a weak basis for explaining the predominance of sex and recombination.

If negative epistasis fails to explain the evolution of sex and recombination, what other process might be responsible? For insight into this question, it is worth returning to some of the earliest arguments for the evolution of sex. Morgan (1913), Fisher (1930), and Muller (1932) argued that an important advantage of sex is that beneficial mutations that arise in separate individuals can be brought together within the same individual. In the absence of sex, only one lineage will leave descendants in the long term, and beneficial mutations in all other asexual lineages will be lost. Thus, the probability of fixation of a new beneficial mutation is severely limited in the absence of sex by the fate of the rest of the genome in which it arises.

The source of disequilibrium implicit in the words of Morgan, Fisher, and Wright is not epistasis but random sampling. When a beneficial allele first arises, it occurs on ("samples") one particular genetic background and, unless recombination occurs, the fate of the beneficial allele is tied to the evolutionary fate of that background. If the population were infinitely large, this sampling problem would disappear because all possible combinations of alleles would be generated instantaneously by mutation. Simulations of finite populations by Hill and Robertson (1966) demonstrated that the probability of fixation of beneficial alleles is limited by linkage among selected loci, a phenomenon now known as the Hill-Robertson effect (see also Felsenstein 1974). Furthermore, both Hill and Robertson (1966) and Felsenstein (1974) demonstrated that, even when epistasis is absent, the disequilibria that develop in a finite population tend to slow the spread of beneficial alleles, especially when recombination is rare. That is, even when fitness is multiplicative across loci, drift generates negative disequilibria among selected loci such that beneficial alleles at one locus become associated with deleterious alleles at other loci. In subsequent simulations, Felsenstein and Yokoyama (1976) found that modifiers that increase the rate of recombination spread in response to the negative disequilibria created by drift in the presence of selection.

The Hill-Robertson effect provides an alternative source for the negative linkage disequilibria that cause indirect selection for modifiers that increase recombination. The way in which random drift generates negative associations is most easily seen by considering two beneficial mutations that arise at different loci in the same generation. Let D measure the gametic linkage disequilibrium between the two loci in a diploid population of size N. Initially, the expected linkage disequilibrium is zero, because the very high chance that the alleles arise on different chromosomes [$D = -1/(2N)^2$ with probability $1 - 1/(2N)$] is counterbalanced by the extremely small chance of a strong positive association if the mutations arise in coupling [$D = (1 - 1/(2N))1/(2N)$ with probability $1/(2N)$]. Consider what would happen to the disequilibrium in the absence of recombination assuming that the mutations rise deterministically in frequency after their appearance. At first, both the negative and the positive disequilibria grow exponentially as the beneficial alleles increase within the population. Eventually, however, the growth of the positive disequilibrium slows down as the coupled beneficial alleles reach fixation. That is, beneficial alleles that arise in coupling rise rapidly and fix within the population, leading to the disappearance of the positive disequilibrium. The negative disequilibrium persists for a much longer period of time, until one or the other allele becomes fixed. Therefore, with multiplicative selection, the variance in disequilibrium present in the first generation ultimately leads to negative disequilibrium, on average.

This heuristic argument applies not only when dis-

equilibrium is generated by the random appearance of mutations on particular genetic backgrounds, but also more generally when disequilibrium is generated by random genetic drift (Qureshi 1963; Latter 1965; Hill and Robertson 1966; Felsenstein 1974). As shown by Hill and Robertson, negative disequilibrium accumulates, on average, whenever genetic drift occurs in the presence of multiplicative selection. Although drift can generate both positive and negative disequilibria, which cancel out on average, selection does not act symmetrically upon the disequilibria thus created. Negative associations persist for longer simply because such associations ("good alleles on bad genetic backgrounds") impede selection. Conversely, positive associations are relatively transient, as selection fixes the best combinations (good with good) and eliminates the worst ones (bad with bad). When averaged across replicates or loci or time, the overall level of disequilibrium expected between selected loci becomes more negative than expected in an infinite population. As we shall see, this process can be significant even in fairly large populations as long as the allele frequencies are initially low and subject to drift.

Analyzing the Hill-Robertson effect is made difficult by the fact that it requires a stochastic model tracking the frequencies of several chromosomes, each of which changes over time in a nonlinear fashion. While difficult, incorporating stochasticity into models of the evolution of sex and recombination is essential, because the fate of a modifier allele depends critically on whether or not drift is included (Felsenstein and Yokoyama 1976; Otto and Barton 1997, 2001). Previous analytical results have used branching processes with multiplicative selection to examine the probability that beneficial alleles are lost or fixed from very large populations. Barton (1995b) calculated the amount by which recombination increases the probability of fixation of beneficial alleles. We extended this analysis to determine the extent to which increased recombination would be favored at a modifier locus (Otto and Barton 1997). Because beneficial mutations are more likely to fix when associated with modifier alleles that increase recombination, these modifier alleles get dragged along as the beneficial alleles spread through the population. This process of genetic hitchhiking causes the modifier alleles, and hence the recombination rate, to increase within the population.

These analyses focused only on the linkage disequilibrium generated by the random occurrence of mutations on specific genetic backgrounds. Even when beneficial alleles are sufficiently common that their fixation is assured, however, random fluctuations in genotype frequencies generate linkage disequilibria among selected loci. Indeed, in our computer simulations, linkage disequilibria generated among beneficial mutations after they first arose affected the dynamics of a modifier of recombination (Otto and Barton 1997). Thus, we must consider the dynamics of disequilibrium generated by random genetic drift over the entire time course as beneficial alleles spread through a population. Because drift tends to generate negative linkage disequilibria, on average (with good alleles residing on bad genetic backgrounds), modifier alleles that increase the rate of recombination bring together good alleles from different individuals, generating indirect selection for recombination. We showed through simulation that this indirect selection is often stronger than the indirect selection for recombination generated by epistasis in populations of small ($2N = 100$) to intermediate ($2N = 10{,}000$) size (Otto and Barton 2001). More importantly, modifiers that increased the frequency of sex and recombination spread, on average, across a broad range of epistasis, from weak negative epistasis to positive epistasis, when drift was incorporated.

While previous analyses have shown that fluctuations in linkage disequilibrium can substantially increase the effects of a favorable allele on linked neutral loci (Stephan *et al.* 1992; Barton 1998), we lack an analytical framework to predict the dynamics of disequilibrium between selected loci in the presence of drift and selection. This article develops such a framework, allowing us to track fluctuations in linkage disequilibrium during the spread of beneficial alleles and to measure the impact on a modifier allele of recombination.

Our analysis makes two crucial simplifications: that the population size (N) is large enough that only leading terms ($O(1/N)$) need be included and that no allele is very rare. Furthermore, because we wish to concentrate on stochastic effects, we assume multiplicative fitnesses throughout. Including epistasis would induce linkage disequilibria and selection on recombination even in an infinite population. These simplifications allow us to approximate the full stochastic dynamics with recursion equations giving the mean and variance for the allele frequencies and disequilibria over time. Approximate solutions to these recursions are then obtained under the assumption of weak selection. This analysis shows that genetic drift creates variance in linkage disequilibrium, which in turn generates negative linkage disequilibrium, on average. The amount of linkage disequilibrium is larger when the beneficial alleles start at low frequency, when selection favoring the alleles is strong, and when linkage is tight. Under these circumstances, a small amount of linkage disequilibrium generated by random genetic drift while the alleles are rare is amplified by selection and only slowly destroyed by recombination. Finally, we determine the amount of indirect selection for recombination generated by drift in the presence of selection by examining evolutionary changes at a third locus that modifies recombination rates. By destroying the negative disequilibrium built up by drift in the presence of directional selection, a modifier allele that increases recombination rates im-

proves the response of its carriers to selection and rises in frequency as a consequence.

TWO-LOCUS MODEL

We begin by examining the linkage disequilibrium generated in a model with two loci, j and k, separated by a recombination rate of r_{jk}. We count genotypic frequencies immediately after meiosis and assume that multiplicative viability selection acts within the population. At locus j, alleles Q_j, P_j segregate at frequencies q_j and p_j and are assumed to have relative fitnesses $(1 - s_j/2):(1 + s_j/2)$, respectively. Similarly, allele P_k has a selective advantage of s_k over allele Q_k at locus k. This assumes that either the population is haploid or fitnesses in a diploid population are multiplicative within as well as between loci [*i.e.*, *QQ:PQ:PP* have fitnesses $(1 - s/2)^2$: $(1 - s^2/4):(1 + s/2)^2$]. This notation is consistent with that of Barton and Turelli (1991) and Barton (1995a) and simplifies the algebra.

In an infinite population, if gametic linkage disequilibrium ($D_{jk} = \text{freq}(P_jP_k)\text{freq}(Q_jQ_k) - \text{freq}(P_jQ_k)\text{freq}(Q_jP_k)$) is initially absent, it remains zero under multiplicative selection (Maynard Smith 1968). In this case, the allele frequencies at each locus increase logistically. Specifically, the ratio p/q increases by a factor $(1 + s/2)/(1 - s/2)$ every generation or by approximately $\exp(st)$ after t generations. The technique that we develop in this article assumes that the trajectory of allele frequencies closely follows the deterministic equations for this model but that small perturbations from this trajectory are caused by random drift. We develop recursion equations for the first and second moments of these perturbations to determine the combined effects of drift and selection over time.

Fluctuations around the deterministic trajectory: We assume a discrete-generation model in which a population is chosen by random sampling from the propagules produced by the previous generation. The frequencies of the chromosomes among the propagules are given by the deterministic recursions for either a haploid population (with selection on haploids, followed by random mating and meiosis, to produce haploid propagules) or a diploid population (with selection on diploids, followed by meiosis and random mating to produce diploid propagules). From these propagules, either $2N$ haploid or N diploid juveniles are sampled with replacement, implying that the chromosome frequencies follow a multinomial distribution. We census the population immediately after sampling occurs.

Random sampling in any particular generation causes genotype frequencies to differ from that expected on the basis of the propagule frequencies. We measure the perturbations generated by a single round of sampling by the random variables ζ_j, ζ_k (for perturbations in the allele frequencies at loci j, k) and ζ_{jk} (for perturbations in the linkage disequilibrium). The exact equations for change in the allele frequencies and linkage disequilibrium due to selection, recombination, and random genetic drift are given in appendix a.

The moments of the multinomial distribution can be used to determine the expected values of the perturbations generated by a single round of random sampling, as well as their variances and covariances. For instance, because random genetic drift does not cause any directional change in allele frequency, $E[\zeta_j] = 0$, but the allele frequency will vary around its deterministic trajectory by an amount that is inversely proportional to the population size with $\text{Var}[\zeta_j] = p_jq_j/(2N)$. The first and second moments for the perturbations are given in appendix a. Higher-order moments are of the order $1/N^2$ or smaller and are ignored in our analysis.

To describe the dynamics over multiple generations, we must keep track of the cumulative deviations from the deterministic trajectory, which we do using the vector $\mathbf{z} = (\delta p_j, \delta p_k, \delta D_{jk})$. The random variable δp_j measures the difference between the actual allele frequency at any point in time and the allele frequency predicted in the absence of random genetic drift. Similarly, δD_{jk} measures the difference between the actual linkage disequilibrium and the disequilibrium predicted deterministically. Throughout, we assume that the perturbations in $\mathbf{z}$ remain small and keep only the leading-order terms.

In the text, we assume that there is no epistasis and that the initial level of disequilibrium, if present, is of the order N^{-1}. Under these assumptions, the disequilibrium predicted deterministically can be set to zero, so that δD_{jk} measures the actual amount of disequilibrium in the presence of drift. The method can be extended, however, to account for epistasis or for strong initial disequilibrium by keeping track of perturbations from the deterministic trajectory for D_{jk} (see appendix a).

Recursions for cumulative effects of drift in the presence of selection can be written as $\mathbf{z}^* = f(\mathbf{z})$ and can be determined from (A2). Assuming that the perturbations are small, $\mathbf{z}^*$ can be expanded in a Taylor series,

$$z_a^* = \left(\sum_b \frac{\partial f_a}{\partial z_b} z_b + \frac{1}{2}\sum_{b,c} \frac{\partial^2 f_a}{\partial z_b \partial z_c} z_b z_c + \cdots\right) + \zeta_a, \qquad (1)$$

where the subscripts (a, b, c) are set to j or k when referring to deviations in the allele frequencies and to jk when referring to the disequilibrium. The partial derivatives in (1) measure the sensitivity of each recursion to the perturbations caused by random sampling. They were derived from the recursions (A2) using the computer algebra package *Mathematica* (Wolfram 1991) and are given in supplementary information (S1; http://www.genetics.org/supplemental/). Note that all partial derivatives are evaluated along the deterministic trajectory ($\mathbf{z} = \mathbf{0}$) and are therefore independent of the perturbations. Taking the expectation of (1) gives

$$E[z_a^*] = \left(\sum_b \frac{\partial f_a}{\partial z_b} E[z_b] + \frac{1}{2}\sum_{b,c} \frac{\partial^2 f_a}{\partial z_b \partial z_c} E[z_b z_c] + \cdots\right) + E[\zeta_a]. \qquad (2)$$

Similarly, second moments are given by the expectation of the product of (1) with itself:

$$E[z_a^* z_b^*] = \sum_{c,d} \frac{\partial f_a}{\partial z_c}\frac{\partial f_b}{\partial z_d} E[z_c z_d] + \sum_c \frac{\partial f_a}{\partial z_c} E[z_c \zeta_b] + \sum_c \frac{\partial f_b}{\partial z_c} E[z_c \zeta_a] + E[\zeta_a \zeta_b] + \cdots. \tag{3}$$

Contributions to the first- and second-order moments arise by drift and are of $O(N^{-1})$ as long as the allele frequencies are not too small (appendix a). In contrast, starting from a population with a specific genotypic composition, higher moments are initially absent, are generated by drift only to $O(N^{-2})$, and are henceforth ignored. Note that the recursion equation (3) for $E[z_a^* z_b^*]$ also describes the change in the covariance, $\mathrm{Cov}[z_a^*, z_b^*] = E[z_a^* z_b^*] - E[z_a^*]E[z_b^*]$, because $E[z_a^*]E[z_b^*]$ is $O(N^{-2})$ and can be ignored. Similarly, $E[(z_a^*)^2]$ describes the change in the variance, $\mathrm{Var}[z_a^*]$, of the perturbations around the deterministic trajectory. Finally, the $E[z_c\zeta_a]$ in (3) are also $O(N^{-2})$ and are dropped.

Having dropped terms of $O(N^{-2})$ in (2) and (3) and using the fact that many of the partial derivatives in supplementary information (S1) equal zero, fluctuations in the allele frequency p_j around the deterministic trajectory satisfy the following recursions:

$$\begin{aligned} E[\delta p_j^*] = {} & \frac{\partial \delta p_j^*}{\partial \delta p_j} E[\delta p_j] + \frac{\partial \delta p_j^*}{\partial \delta D_{jk}} E[\delta D_{jk}] + \frac{1}{2}\frac{\partial^2 \delta p_j^*}{\partial \delta p_j^2} E[\delta p_j^2] \\ & + \frac{1}{2}\frac{\partial^2 \delta p_j^*}{\partial \delta D_{jk}^2} E[\delta D_{jk}^2] + \frac{\partial^2 \delta p_j^*}{\partial \delta p_j \partial \delta D_{jk}} E[\delta p_j \delta D_{jk}] \\ & + \frac{\partial^2 \delta p_j^*}{\partial \delta p_k \partial \delta D_{jk}} E[\delta p_k \delta D_{jk}] + O(N^{-2}) \end{aligned} \tag{4a}$$

$$\begin{aligned} E[\delta p_j^{*2}] = {} & \left(\frac{\partial \delta p_j^*}{\partial \delta p_j}\right)^2 E[\delta p_j^2] + \left(\frac{\partial \delta p_j^*}{\partial \delta D_{jk}}\right)^2 E[\delta D_{jk}^2] \\ & + 2\left(\frac{\partial \delta p_j^*}{\partial \delta p_j}\right)\left(\frac{\partial \delta p_j^*}{\partial \delta D_{jk}}\right) E[\delta p_j \delta D_{jk}] + \frac{p_j q_j}{2N} + O(N^{-2}). \end{aligned} \tag{4b}$$

Fluctuations in the allele frequency p_k are described by analogous equations, with subscripts j and k interchanged. Of greater relevance to the evolution of recombination are fluctuations that occur in the disequilibrium, δD_{jk}, around the deterministic equilibrium of $D_{jk} = 0$:

$$\begin{aligned} E[\delta D_{jk}^*] = {} & \left(\frac{\partial \delta D_{jk}^*}{\partial \delta D_{jk}}\right) E[\delta D_{jk}] \\ & + \left(\frac{1}{2}\frac{\partial^2 \delta D_{jk}^*}{\partial \delta D_{jk}^2} E[\delta D_{jk}^2] + \frac{\partial^2 \delta D_{jk}^*}{\partial \delta p_j \partial \delta D_{jk}} E[\delta p_j \delta D_{jk}] + \frac{\partial^2 \delta D_{jk}^*}{\partial \delta p_k \partial \delta D_{jk}} E[\delta p_k \delta D_{jk}]\right) \\ & + O(N^{-2}) \end{aligned} \tag{5a}$$

$$E[\delta D_{jk}^{*2}] = \left(\frac{\partial \delta D_{jk}^*}{\partial \delta D_{jk}}\right)^2 E[\delta D_{jk}^2] + \frac{p_j q_j p_k q_k}{2N} + O(N^{-2}) \tag{5b}$$

$$\begin{aligned} E[\delta p_j^* \delta D_{jk}^*] = {} & \frac{\partial \delta D_{jk}^*}{\partial \delta D_{jk}}\left(\frac{\partial \delta p_j^*}{\partial \delta D_{jk}} E[\delta D_{jk}^2] + \frac{\partial \delta p_j^*}{\partial \delta p_j} E[\delta p_j \delta D_{jk}]\right) \\ & + O(N^{-2}). \end{aligned} \tag{5c}$$

$E[\delta p_k^* \delta D_{jk}^*]$ is given by (5c), with subscripts j and k interchanged. These equations apply even if selection coefficients vary through time, provided that there is no frequency dependence. Frequency-dependent selection could be modeled in a similar fashion, taking into account that the partial derivatives will contain additional components depending on the form of selection. With either constant or variable selection coefficients, recursions (4) and (5) can be evaluated numerically (*Mathematica* package available upon request) to predict the means, variances, and covariances of the perturbations over time.

The recursions for the perturbations can be further analyzed under the assumption that selection is weak (appendix b). Solution (B3) for the perturbations involving the disequilibrium is derived assuming that linkage is sufficiently loose relative to selection that the terms involving disequilibria reach a steady-state level over a faster timescale than the changes in allele frequencies; this steady-state level is known as the quasi-linkage equilibrium or QLE (Kimura 1965; Nagylaki 1993; Barton 1995a). Solution (B5) is derived assuming that both selection coefficients and recombination rates are small enough that the recursions can be well approximated by differential equations, which are then solved.

Development of disequilibria under directional selection: We now consider how the above equations can be used to infer how disequilibrium develops in finite populations. Variance in linkage disequilibrium is generated directly by random drift, contributing the $p_j q_j p_k q_k / 2N$ term to (5b). As long as the direction of selection remains the same, all of the terms in (5c) are positive ($\partial \delta p_j^* / \partial \delta D_{jk}$, $\partial \delta p_k^* / \partial \delta D_{jk}$, $\partial \delta D_{jk}^* / \partial \delta D_{jk} > 0$), indicating that a positive covariance between the frequency of each beneficial allele and linkage disequilibrium will develop as a result of this variance in linkage disequilibrium. This covariance arises because allele frequencies increase more rapidly in populations that happen to have positive linkage disequilibrium. The presence of both variance in disequilibrium and covariance between allele frequencies and linkage disequilibrium then causes the expected linkage disequilibrium (5a) to depart from zero. Because $\partial^2 \delta D_{jk}^* / \partial \delta D_{jk}^2$ is negative (appendix a), variance in disequilibrium across replicate populations leads, on average, to negative disequilibrium. Mathematically, this occurs because the mean fitness, $\overline{W}$, given by (A1) is higher with positive disequilibrium than with negative disequilibrium for any given set of allele frequencies. Consequently, the amount of positive disequilibrium gets divided by a larger quantity ($\overline{W}^2$) in (A2b) and becomes relatively smaller in the next generation. Another, perhaps more intuitive, explanation, is that the expected disequilibrium across replicates becomes negative because selection is more efficient when good alleles are found on good genetic backgrounds (*i.e.*, when disequilibrium is positive by chance), causing positive disequilibrium to decay more rapidly than negative disequilibrium whenever it occurs. The positive covariance generated between each allele

frequency and the linkage disequilibrium leads the expected disequilibrium to become even more negative ($\partial^2 \delta D^*_{jk}/\partial \delta p_j \partial \delta D_{jk}$, $\partial^2 D^*_{jk}/\partial \delta p_k \partial \delta D_{jk} < 0$). This occurs because those populations that happen to have positive disequilibrium and, consequently, faster rising allele frequencies (generating the positive covariance) also have higher mean fitness values in future generations. The linkage disequilibrium in future generations will then continue to be divided by a larger quantity ($\overline{W}^2$) in (A2b), compounding the reduction in positive disequilibrium relative to negative disequilibrium. As a consequence of all of these effects, the expected value of linkage disequilibrium becomes negative even though it was initially zero and even though selection acts independently on the two loci.

Provided that allele frequencies are intermediate [*i.e.*, p, q are $O(1)$], the average amount of negative disequilibrium that develops is always proportional to $1/N$: if the population size is doubled, the amount of disequilibrium generated by drift and selection is halved. When linkage is loose, disequilibrium does not accumulate appreciably over time. With tight linkage, however, substantial amounts of disequilibrium can accumulate within a population, especially when allele frequencies are initially low. This is a consequence of the fact that a small amount of disequilibrium generated by drift between rare alleles becomes a large amount of disequilibrium as the alleles approach a frequency of ½. If the initial allele frequency is too small [$O(1/N)$], as is the case with a single favorable mutation, then higher-order terms in the Taylor series approximation (1) become appreciable and can no longer be ignored.

Figure 1 shows the distribution of allele frequencies over time in a population of size $2N = 10{,}000$ when alleles P_j and P_k are initially at frequency $p_0 = 0.01$ ($s_j = s_k = r_{jk} = 0.1$). The thick curve shows the mean trajectory (which is nearly equal to the deterministic trajectory), while the thin curves indicate the range within which allele frequencies lie 95% of the time as calculated from (4). These predictions were compared to simulation results based on a Monte Carlo simulation that sampled $2N$ chromosomes each generation using a multinomial distribution, with parameters equal to the frequency of each chromosome given by the deterministic recursions. Simulations (dots) indicate that (4) accurately captures the effects of drift on allele frequencies. Note that drift causes substantial variation in the trajectory of allele frequencies even in a population of size 10,000. This variation arises primarily during the early stages when only a small number of beneficial alleles are present ($2Np_0 = 100$) and represents random accelerations or decelerations in the spread of favorable alleles. On average, beneficial alleles spread slightly less rapidly than expected in an infinite population, indicating that the Hill-Robertson effect is present, if small, in populations of size 10,000. The effects of drift become even more important when there are initially fewer alleles within the population (*e.g.*, $2Np_0 \leq 10$). In this case, however, iterating (4) and (5) overestimates the Hill-Robertson effect (results not shown), because higher-order moments in the perturbations are ignored and yet play an important role whenever allele frequencies are near zero.

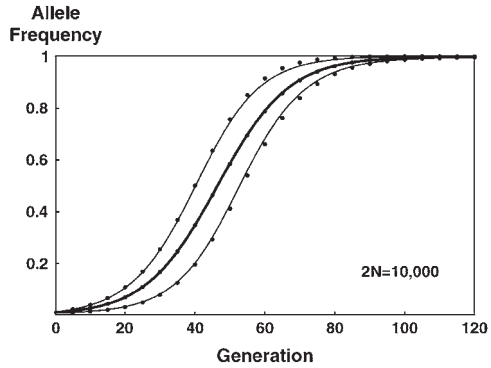

FIGURE 1.—Allele frequency change in a finite population. The thick curve shows the average frequency over time of the allele P_j, with selective advantage $s_j = 0.1$ and initial frequency $p_{j0} = 0.01$, in a population of size $2N = 10{,}000$, using Equation 4a to determine the departure from the deterministic trajectory ($p_j/q_j = e^{s_j t} p_{j0}/q_{j0}$). Although drift causes P_j to rise less rapidly on average than expected in an infinitely large population, the effect is small, shifting the trajectory to the right by an amount approximately equal to the thickness of the curve. We therefore use the deterministic trajectory to approximate p_j throughout this article, which is a reasonable assumption provided that $Ns_j \gg 1$. The thin curves show the expected allele frequency ±2 standard deviations based on (4b) (a weak selection approximation generates indistinguishable results and is not shown). Dots show simulation results for the mean allele frequency ±2 standard deviations, based on 1,000,000 replicates.

Figure 2 shows how drift affects (a) the variance in disequilibrium scaled relative to the allele frequencies, $E[\delta D^2_{jk}]/(p_j q_j p_k q_k)$, (b) the correlation between δp_j and disequilibrium, $E[\delta p_j \delta D_{jk}]/\sqrt{E[\delta p_j^2] E[\delta D^2_{jk}]}$, and (c) the average disequilibrium. The average magnitude of the linkage disequilibrium is greater when the initial frequency of the beneficial alleles is lower, because random associations generated in the early stages persist and are amplified as the beneficial alleles rise in frequency (compare dotted, dashed, and solid curves, which correspond to initial frequencies 0.001, 0.01, and 0.1). In this example, where $s_j = s_k = r_{jk} = 0.1$, the QLE approximation (B3) fails to capture the influence of the initial allele frequency, which is pronounced when selection is strong relative to recombination [*e.g.*, the QLE prediction (B3a) for the scaled variance is constant at 0.00053]. On the other hand, the weak selection approximation (B5) overestimates the impact of initial fluctuations (compare thin to thick curves), although calculations with weaker selection (0.01) show that the approximation then agrees closely.

Figure 3 shows that the expected disequilibrium scales

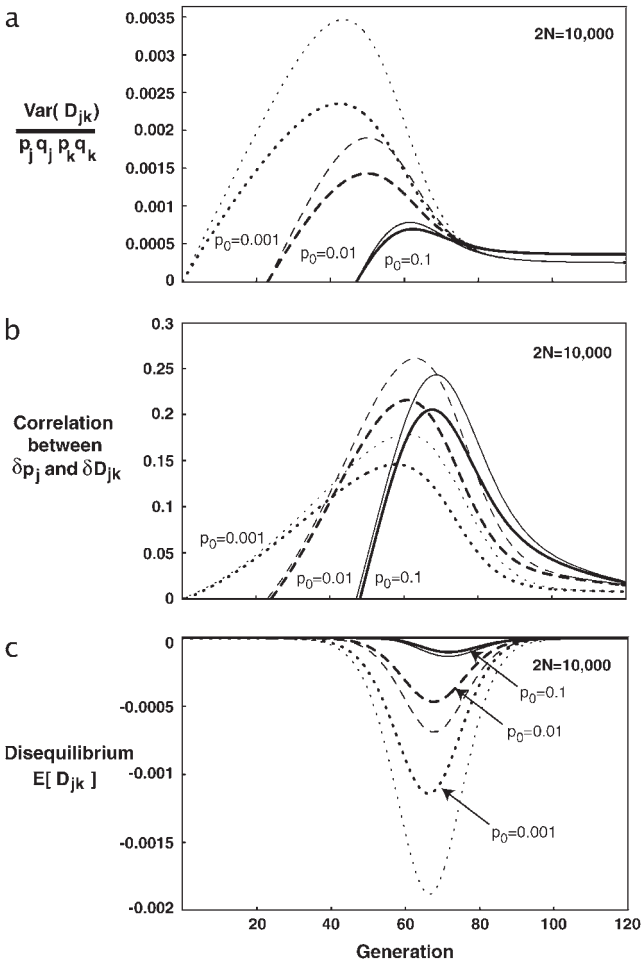

FIGURE 2.—Effects of drift and selection on disequilibrium in a finite population under directional selection. (a) Variance in disequilibrium scaled to $p_j q_j p_k q_k$; (b) correlation between effects of drift on allele frequency and on disequilibrium (Corr$[\delta p_j \delta D_{jk}]$); (c) average disequilibrium over time. In each case, the initial allele frequencies are $p_{j0} = p_{k0} = 0.001$ (dotted curves), 0.01 (dashed curves), or 0.1 (solid curves), where the latter curves are shifted to the right by 23.1 and 47.1 generations, respectively, so that allele frequencies are equivalent for all curves at any point in time. The thick curves give the exact calculations from (5), while the thin curves show the weak selection approximation (B5). Results are shown for $s_j = s_k = 0.1$, $r_{jk} = 0.1$, and $2N = 10{,}000$.

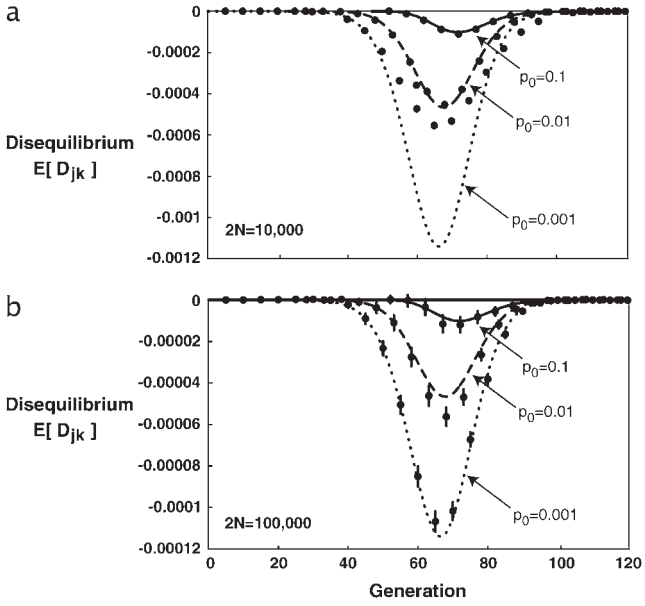

FIGURE 3.—Observed and expected disequilibrium under directional selection. The disequilibria expected from (5) (curves) and observed in simulations with 1,000,000 replicates (dots, with bars indicating ±2 standard errors) are shown for (a) $2N = 10{,}000$ and (b) $2N = 100{,}000$. Equation 5 accurately estimates the amount of disequilibrium generated except when there are only 10 initial copies of each allele ($p_{j0} = p_{k0} = 0.001$ and $2N = 10{,}000$). Otherwise, the disequilibrium scales directly with $1/(2N)$. With only 10 copies initially, both beneficial alleles did not always fix, and we report only the disequilibrium observed in those cases in which both beneficial alleles remain. Remaining parameters and symbols are as in Figure 2.

with $1/(2N)$ and matches the disequilibrium observed in simulations. The only exception is when the initial number of copies is very small (in this example, for $2Np_0 = 10$). With few initial copies of the beneficial alleles, three issues arise that are not accounted for by our analysis. First, the fluctuations around the deterministic trajectory become so large that the Taylor approximation to Equation 1 is no longer adequate. Second, with few initial copies, beneficial alleles do not necessarily fix within the population. Third, those beneficial alleles that do fix are more likely to be positively associated with other beneficial alleles. Therefore, during the first few generations, rare beneficial alleles are more likely to be lost when negative disequilibrium develops (thereby destroying the disequilibrium) and more likely to persist when positive disequilibrium develops. Therefore, the average disequilibrium among beneficial alleles that have survived loss while rare might very well become positive, at least during the initial spread of the beneficial alleles. There is some evidence for this effect within the simulations: in the fifth generation, significantly positive disequilibrium was found when $2N = 10{,}000$ and initial frequencies were 0.001 but in no other case. These positive associations will counteract the accumulation of negative disequilibrium to some extent, explaining, in part, the smaller amount of disequilibrium observed for this case (Figure 3, $p_0 = 0.001$).

Figure 4 shows the most extreme value of the expected linkage disequilibrium observed during the spread of alleles P_j and P_k as a function of r_{jk}. Results from Monte Carlo simulations are well approximated by iterating Equations 4 and 5 (compare dots to thick dashed curve for $p_0 = 0.01$ and thick solid curve for $p_0 = 0.1$). The QLE approximation (B3c; thin curve) is accurate only for loose linkage ($r_{jk} > s_j$, $s_k = 0.1$) and fails to capture the dependence of disequilibrium on initial allele frequencies. With tight linkage, random fluctuations produced when small numbers of the favor-

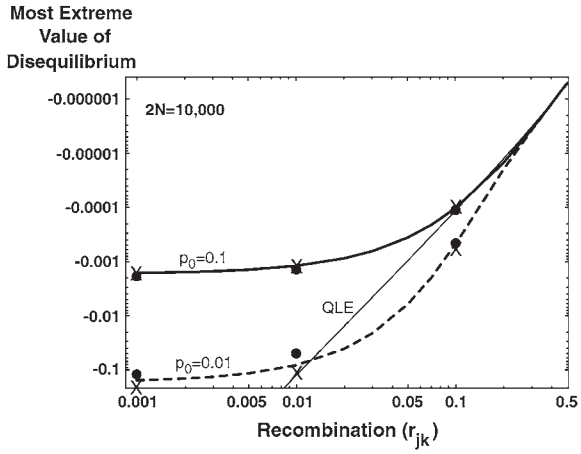

FIGURE 4.—The most extreme value of the average disequilibrium under directional selection, plotted on a log-log scale. During the spread of beneficial alleles, the expected disequilibrium from (5) (thick curves) becomes negative, taking on the minimum values shown. More extreme disequilibria are observed when the initial allele frequencies are low [compare $p_{j0} = p_{k0} = 0.01$ (dashed curves) to $p_{j0} = p_{k0} = 0.1$ (solid curves)] and when recombination is rare, in which case the expected disequilibrium approaches the minimum value that D_{jk} may take (-0.25, corresponding to the bottom axis of the graph). The thin curve gives the most extreme value of the disequilibrium from the QLE approximation (B3c), which is accurate only when $r_{jk} > 0.1$. Dots show the most extreme value of the mean disequilibrium observed in 1,000,000 replicate simulations for $r_{jk} = 0.001$, 0.01, and 0.1. X's represent a numerical evaluation of the weak selection approximation, (B5c). Remaining parameters and symbols are as in Figure 2.

able alleles are present persist to produce large fluctuations when the alleles become common. Consequently, the most extreme disequilibrium observed is very sensitive to starting allele frequencies when linkage is tight. When $p_0 = 0.01$ and $r_{jk} < 0.01$, drift and selection are substantial enough to drive the expected linkage disequilibrium to near its minimum value (-0.25), even though disequilibrium is initially absent and even though selection acts independently on all loci ($s_j = s_k = 0.1$).

Including a modifier of recombination: The negative disequilibrium generated by drift in the presence of directional selection reduces the genotypic variability within a population and slows the response of that population to selection. A modifier allele that increases recombination among the selected genes will regenerate the genetic variation in fitness hidden in linkage disequilibrium. In particular, individuals carrying such a modifier allele are more likely to produce offspring of the fittest genotype. This effect can select for increased recombination at modifier loci, an effect examined in this section. Suppose that a modifier of recombination segregates at locus i. We assume that the modifier has no direct effect on fitness and alters recombination rates by a small amount δr. If terms of $O(\delta r^2)$ are neglected, then the only significant terms involve the change in recombination between the selected loci (denoted $\delta r_{j,k|i}$). The modifier may also alter its own rates of recombination with other loci ($\delta r_{i,j|i}$, etc.) and may show dominance ($\delta r_{j,k|i,i}$). However, these effects are negligible for weak modifiers (BARTON 1995a).

The equations for the effects of drift on the modifier frequency and its associations D_{ij}, D_{ik}, D_{ijk} are given by (A2c–e). Disequilibria are measured as central moments of effects (BARTON 1995a). For two loci, this definition coincides with the standard two-locus measure of gametic disequilibrium. For three loci, D_{ijk} is equal to $\Sigma g[\mathbf{X}](X_i - p_i)(X_j - p_j)(X_k - p_k)$, where the sum is taken over all chromosomes, each at frequency $g[\mathbf{X}]$, where $\mathbf{X}$ is a vector indicating whether or not an allele is present at each locus (*i.e.*, X_i is 0 if allele P_i is absent and 1 if it is present). A modifier allele that increases recombination between loci j and k becomes associated with the fittest genotype, P_jP_k, if that genotype is underrepresented within the population (*i.e.*, if D_{jk} is negative). This association is represented by the first term in (A2e), which generates D_{ijk} in proportion to $-\delta r_{j,k|i} D_{jk}$. This three-way association then generates positive two-way associations, D_{ij} and D_{ik}, between a modifier allele that increases recombination and the beneficial alleles at loci j and k (A2d). All three associations, D_{ij}, D_{ik}, D_{ijk}, cause changes in the frequency of the modifier allele (A2c). Note that because the modifier is assumed to have a small effect, these equations are linear in D_{ij}, D_{ik}, D_{ijk}, and all will change in proportion to $\delta r_{j,k|i}$.

Equations 2 and 3 can be used to find the mean and variance of the three-locus system, as before. The first step is to find the first and second differentials of (A2c–e), such as $\partial\delta p_i^*/\partial\delta D_{jk}$. These differentials are calculated as in supplementary information (S1) and are available upon request. The second step is to calculate the expected values and the covariances for the perturbations involving the modifier by incorporating these differentials into Equations 2 and 3. Following this method and dropping terms of $O(N^{-2})$, the recursion describing the cumulative effects of drift and selection at loci j and k on the frequency of a modifier allele at locus i is

$$\begin{aligned} E[\delta p_i^*] = E[\delta p_i] &+ \frac{s_j}{\phi_j}E[\delta D_{ij}] + \frac{s_k}{\phi_k}E[\delta D_{ik}] + \frac{s_j s_k}{\phi_j\phi_k}E[\delta D_{ijk}] \\ &- (\phi_k^2 s_j^2 E[\delta p_j \delta D_{ij}] + \phi_k s_j^2 s_k E[\delta p_j \delta D_{ijk}] \\ &\quad + \phi_j^2 s_k^2 E[\delta p_k \delta D_{ik}] + \phi_j s_j s_k^2 E[\delta p_k \delta D_{ijk}] \\ &\quad + \phi_k s_j^2 s_k E[\delta D_{jk}\delta D_{ij}] + \phi_j s_j s_k^2 E[\delta D_{jk}\delta D_{ik}] \\ &\quad + s_j^2 s_k^2 E[\delta D_{jk}\delta D_{ijk}])/(\phi_j^2\phi_k^2), \end{aligned} \tag{6}$$

where $\phi_j = 1 + s_j(p_j - q_j)/2$ and $\phi_k = 1 + s_k(p_k - q_k)/2$. The expected change in the modifier is driven both by expectation terms $E[\delta D_{ij}]$ and by covariances such as $E[\delta p_j \delta D_{ij}]$, each of which is proportional to $\delta r_{j,k|i}/(2N)$. The recursions for these expectation terms are compli-

cated but can be calculated symbolically in a straightforward way from (2) and (3). These recursions may be iterated numerically or may be approximated by assuming that selection is weak as described in supplementary information (S2; http://www.genetics.org/supplemental/).

Assuming weak selection and loose linkage, the system approaches quasi-linkage equilibrium, and the per generation change in the frequency of a modifier at QLE can be found (supplementary information, S2). This gives a complicated function of the recombination rates (S2.4), which is proportional to $\delta r_{j,k|i} p_i q_i V_j V_k / 2N$, where $V_j = 2s_j^2 p_j q_j$ and $V_k = 2s_k^2 p_k q_k$ are the additive genetic variances in diploid fitness at the two selected loci. Modifier alleles that increase recombination ($\delta r_{j,k|i} > 0$) thus increase in frequency even in the absence of epistasis in finite populations.

Expression (S2.4) takes a simple form if there is no genetic interference and if recombination rates are small in absolute terms but large relative to the selection coefficients, as required for QLE (*i.e.*, s_j, $s_k \ll r_{jk}$, r_{ij}, r_{ik}, $r_{ijk} \ll 1$). Assuming, without loss of generality, that locus k is the right-most locus, we can describe the relative position of the modifier by defining $\alpha = r_{ik}/r_{jk}$; α varies between zero and one when the gene order is *jik* and is greater than one when the gene order is *ijk*. Then, keeping only the leading-order term in (S2.4), the per generation change in the modifier becomes

$$E[\delta p_i^*] = E[\delta p_i] + \frac{\delta r_{j,k|i} p_i q_i V_j V_k}{8N r_{jk}^5} g(\alpha), \qquad (7a)$$

where

$$g(\alpha) = \frac{8 + 9\alpha - 5\alpha^2 - 8\alpha^3 + 4\alpha^4}{2(1-\alpha)^2\alpha^2(1+\alpha)(2-\alpha)} \quad \text{for } 0 < \alpha < 1 \text{ (gene order } jik) \qquad (7b)$$

$$g(\alpha) = \frac{7 - 5\alpha - 16\alpha^2 + 10\alpha^3 + 8\alpha^4 + 4\alpha^5}{(\alpha^2-1)^2\alpha^3} \quad \text{for } 1 < \alpha \text{ (gene order } ijk). \qquad (7c)$$

Figure 5 illustrates $g(\alpha)$ and shows how evolution of a modifier of recombination depends on the relative positions of the loci, α. The function $g(\alpha)$ rises rapidly near $\alpha = 0$ and $\alpha = 1$, where it is approximately $g(\alpha) \approx 2/\alpha^2$ and $2/(1-\alpha)^2$, respectively, although the QLE approximation breaks down if recombination rates become too small relative to selection. For gene order *jik*, $g(\alpha)$ reaches a minimum value of 37⅓, when the modifier is halfway between the selected loci. For gene order *ijk*, $g(\alpha)$ lies between $2/(1-\alpha)^2$ and $4/(1-\alpha)^2$ and decreases rapidly for relatively distant modifier loci.

Equation 7a demonstrates that the effect that drift and selection have on a modifier of recombination falls precipitously as recombination becomes more frequent within the genome ($\sim r_{jk}^{-5}$). If selected loci are scattered randomly over a genome, then the effect of a modifier on tightly linked loci will dominate its dynamics. As recombination rates become comparable with the selection coefficients, however, the QLE assumptions break down, making it impossible to determine the average force on the modifier when selected loci are scattered randomly across a genome. Note that as selection and recombination decrease together toward zero in (7a), selection on the modifier approaches a positive limit, inversely proportional to population size (*i.e.*, if s, r, and δr are all small and of the same order, $\Delta E[\delta p_i] \sim 1/N$). This limit appears puzzling, because the average frequency of the modifier should not change in the absence of selection. There is no contradiction here, because we have assumed throughout that $s \gg 1/N$ when dropping terms of $O(N^{-2})$. Thus, our approximations leading to (7) require that N increases toward infinity as s decreases toward zero, causing $\Delta E[\delta p_i]$ to decrease implicitly as selection weakens.

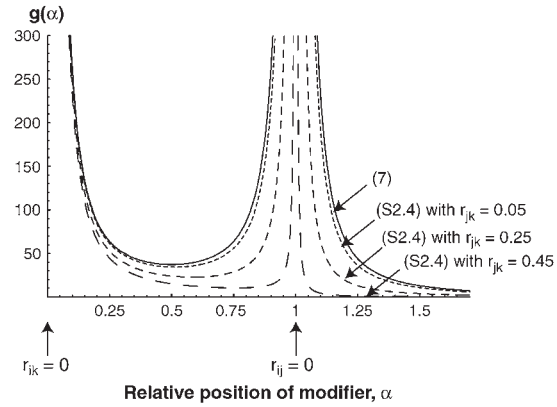

FIGURE 5.—The influence of the relative position of the modifier, $\alpha = r_{ik}/r_{jk}$, on $g(\alpha)$. According to the QLE approximation (7), the change in frequency of a modifier is proportional to $g(\alpha)$ (solid curve) when recombination rates are small but larger than the selection coefficients. The dashed curve compares this approximation to the full QLE result [the complicated function $f(r)$ in (S2.4) with no interference], which applies even when recombination rates are high.

In Figures 6–9, we describe changes in the frequency of the modifier using the standardized function, $\beta_i = \Delta p_i/(\delta r_{j,k|i} p_i q_i)$, which we call the *selection gradient* acting on the modifier allele. When multiplied by the genetic variance in recombination within the population, $V_i = 2\delta r_{j,k|i}^2 p_i q_i$, the selection gradient describes the expected increase in recombination ($= 2\Delta p_i \delta r_{j,k|i}$). The selection gradient can also be related to the effective selection coefficient, s_e, defined as the function that causes Δp_i to equal $s_e p_i q_i$. For a modifier of effect $\delta r_{j,k|i}$, the effective selection coefficient is $s_e = \beta_i \delta r_{j,k|i}$. It is important to recall that the modifier does not directly experience selection. Rather, the selection gradient measures the indirect selection arising from genetic associations between the modifier locus and the directly selected loci j and k. To describe the total effect of selection at linked loci on recombination, we use the *cumulative selection gradient*, $\Sigma\beta_i$, which represents the sum of β_i from the initial generation up until a given generation, or the

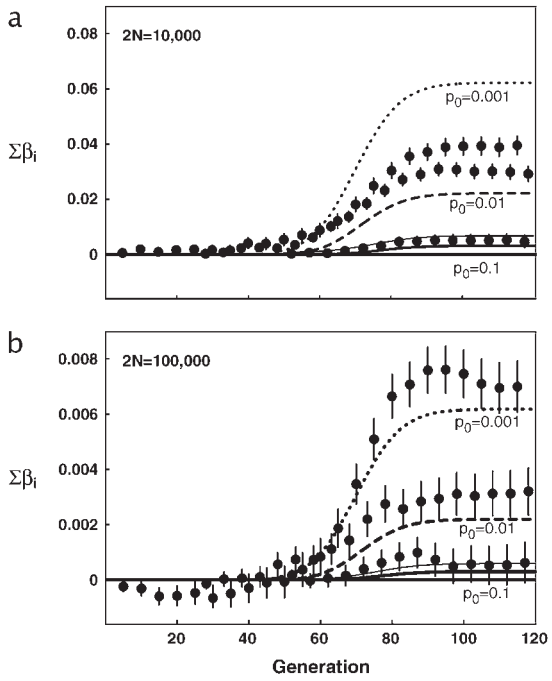

FIGURE 6.—The cumulative selection gradient, $\Sigma\beta_i$, on a modifier of recombination as a function of time under directional selection for (a) $2N = 10{,}000$ and (b) $2N = 100{,}000$. The selection gradient per generation, β_i, is given by the rise in these curves over a single generation. Equations 2 and 3 are used to generate the curves for $p_{j0} = p_{k0} = 0.001$ (dotted curves), 0.01 (dashed curves), and 0.1 (thick solid curves), where the latter two are shifted to the right as in Figure 2. Simulations (dots) are based on 10,000,000 replicates for a modifier that changes recombination by $\delta r_{j,k|i} = 0.05$ and that starts at frequency $p_{i0} = 0.5$. With $p_{j0} = p_{k0} = 0.001$ and $2N = 10{,}000$, both alleles did not always fix, and we report changes in the modifier frequency conditional on fixation. The QLE prediction (S2.4) is shown as a thin solid curve and fails to account for the sensitivity to initial allele frequencies when recombination is not large relative to selection. Remaining parameters are $s_j = s_k = 0.1$ and $r_{ij} = r_{jk} = 0.1$, with gene order *ijk* and no crossover interference.

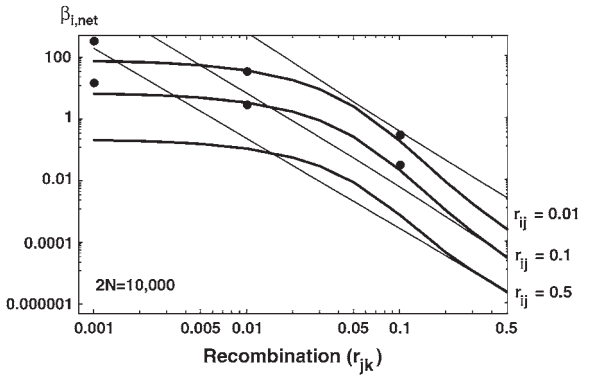

FIGURE 7.—The net selection gradient, $\beta_{i,\text{net}}$, on a modifier of recombination as a function of the recombination rates under directional selection, plotted on a log-log scale. Equations 2 and 3 are used to generate the thick curves for $r_{ij} = 0.01$, 0.1, and 0.5 (from top to bottom). The thin curves give the net change in modifier frequency using the QLE approximation (S2.4) for the same range of r_{ij}. Dots show simulation results based on 1,000,000 replicates for $r_{jk} = 0.001$, 0.01, and 0.1 and $r_{ij} = 0.01$ and 0.1 (from top to bottom; standard errors were too large for $r_{ij} = 0.5$ to assess the effect). In the simulations, the modifier changes recombination by an amount $\delta r_{j,k|i} = r_{jk}/2$ and starts at frequency $p_{i0} = 0.5$. The gene order is *ijk*, and there is no interference. Remaining parameters are $s_j = s_k = 0.1$, $p_{j0} = p_{k0} = 0.01$, and $2N = 10{,}000$, with gene order *ijk* and no crossover interference.

net selection gradient, $\beta_{i,\text{net}}$, which represents the sum of β_i over the course of selection at loci *j* and *k* until fixation of the beneficial alleles.

As a consequence of the fact that disequilibrium is more likely to develop by drift when allele frequencies are initially low at selected loci (Figure 3), the selection gradient acting on a modifier is larger under these conditions, as shown in Figure 6. Simulation results for the change in frequency of the modifier are in general agreement with iteration of (2) and (3), even though the assumptions of weak selection and a weak modifier are violated in the simulations (where $s_j = s_k = 0.1$ and $\delta r_{j,k|i} = r_{jk}/2$). The largest discrepancies are observed when the initial number of alleles is low (*e.g.*, when $2N = 10{,}000$ and $p_0 = 0.001$), such that the perturbations due to drift become larger than allowed by our analysis. It is worth emphasizing that the simulation results presented in Figure 6 represent the mean ± standard errors based on 10^6 replicates. Consequently, the standard deviations are enormous (multiply the length of the bars by 1000). Thus, while our analysis accurately predicts the mean change in frequency of a modifier, the effect of one bout of selection on a modifier is highly variable with only two selected loci and one modifier locus. Only when many loci affect fitness and recombination does the frequency of a modifier of recombination increase consistently (OTTO and BARTON 2001; ILES *et al.* 2003).

The net selection gradient, $\beta_{i,\text{net}}$, acting on a modifier over the entire time course of substitution at two linked loci is shown in Figure 7. The net increase of the modifier can be large with tight linkage: for example, with $r_{ij} = r_{jk} = 0.01$, $\Delta p_i = 38\ p_i q_i \delta r_{j,k|i}$. For very tight linkage, the change in the modifier is underestimated by iteration of (2) and (3) because the disequilibria caused by drift become substantial and higher-order effects begin to contribute [*e.g.*, for $r_{jk} = 0.001$ and $r_{ij} = 0.01$, $\Delta p_i = 334\ p_i q_i \delta r_{j,k|i}$ from simulations but only $75\ p_i q_i \delta r_{j,k|i}$ from (2) and (3)]. The QLE approximation (S2.4) is fairly accurate for loose linkage (r_{ij}, $r_{jk} > s_j$, $s_k = 0.1$) but substantially overestimates the change in modifier frequency for tight linkage, as shown by the thin lines in Figure 7.

Fluctuating polymorphisms: The analysis described in this article can be used whether selection coefficients are constant or vary over time, as long as the population size is reasonably large and the numbers of each allele are never very small. With directional selection, indirect selection on a modifier ceases with the fixation of bene-

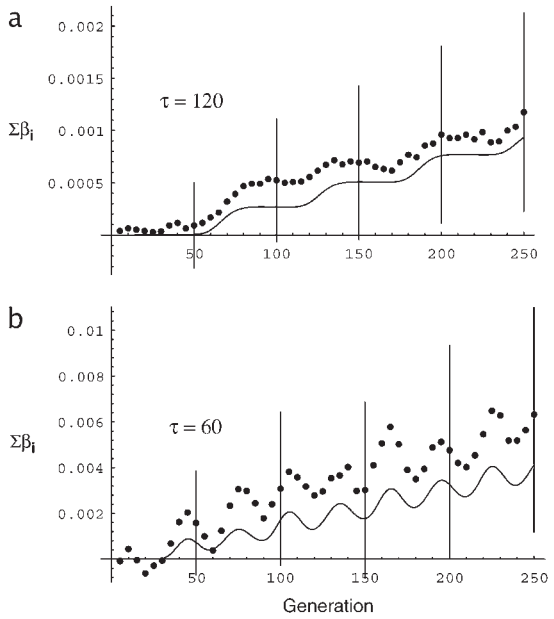

FIGURE 8.—The cumulative selection gradient, $\Sigma\beta_i$, on a modifier of recombination under fluctuating selection. The strength of selection on loci j and k varies sinusoidally over time, with the maximum strength of selection set to $\alpha_j = \alpha_k = 0.1$ and with both loci in phase. The gene order is *ijk*, there is no interference, and $\delta r_{j,k|i} = r_{jk}/2$. Equations 2 and 3 are used to generate the expected change in the modifier (solid curves) starting with $p_{i0} = p_{j0} = p_{k0} = 0.5$ in a population of size $2N = 100{,}000$. Dots show simulation results based on 20,000,000 replicates (± 2 standard errors). (a) Period (τ) is 120 generations, $r_{jk} = 0.1$, and $r_{ij} = 0.1$. (b) Period (τ) is 60 generations, $r_{jk} = 0.02$, and $r_{ij} = 0.01$.

ficial alleles. With fluctuating selection, on the other hand, polymorphism at the selected loci can be maintained over longer periods of time, prolonging selection for increased recombination. Even in the absence of epistatic interactions among loci and even when a population is initially in linkage equilibrium, drift in a finite population subject to fluctuating selection will lead to the accumulation of disequilibrium that reduces genetic variability in fitness. We have used Equations 2 and 3 to track the change in modifier frequency in a population subject to sinusoidal fluctuations in selection at two loci with $s_j[t] = \alpha_j \mathrm{Cos}(2\pi \cdot t/\tau)$ and $s_k[t] = \alpha_k \mathrm{Cos}(2\pi \cdot t/\tau + \Phi)$, where τ is the period (assumed to be the same for both loci) and Φ measures the shift in phase of selection at the two loci (Figure 8). Note that when the direction of selection changes, the selection gradient on the modifier weakens (Figure 8a) or becomes negative (Figure 8b). Nevertheless, recombination rates rise over successive cycles.

The QLE result (S2.4) continues to apply when selection fluctuates over time, as long as recombination is frequent relative to the strength of selection and to the changes over time in selection ($s_j[t]$, $s_k[t]$, $1/\tau \ll r_{ij}$, r_{jk}, r_{ik}, r_{ijk}). Under sinusoidal selection, the product of the additive genetic variances in fitness at the two loci (V_jV_k) is $\sim 3/2(\alpha_j^2 p_j q_j)(\alpha_k^2 p_k q_k)$ if the allele frequencies do not change substantially over a cycle. This result indicates that, when linkage is loose, indirect selection on a modifier will be stronger when the allele frequencies remain near $1/2$, such that the genetic variance in fitness remains high.

For weak selection and very tight linkage ($r\tau \ll 1$), we can obtain an approximate solution to the recursion equations (S2.2) and (S2.3). Assuming that several cycles of fluctuating selection have passed such that the dynamical system has reached a steady state, the solution to these recursions may be approximated as described in supplementary information (S3; http://www.genetics.org/supplemental/). The change in the modifier per generation averaged over one cycle is, to leading order in the recombination rates,

$$E[\delta p_i^*] = E[\delta p_i] + \frac{\delta r_{jk|i} p_i q_i}{2N} \frac{\mathrm{Var}[p_j'] \mathrm{Var}[p_k'](1/r_{ij} + 1/r_{ik}) + \mathrm{Cov}[p_j', p_k']^2 (2/r_{jk} + 6/r_{ijk})}{H_{jk}(2r_{jk}(r_{ijk} + r_{jk}))}, \quad (8)$$

where H_{jk} is the harmonic mean value of $p_j q_j p_k q_k$, $\mathrm{Var}[p_j']$ and $\mathrm{Var}[p_k']$ measure the variance in the rate of allele frequency change over the cycle, and $\mathrm{Cov}[p_j', p_k']$ measures the covariance in these rates at the two loci:

$$\mathrm{Var}[p_j'] = \frac{\sum_{j=0}^{\tau-1} (s_j p_j q_j)^2}{\tau} = \frac{\sum_{j=0}^{\tau-1} (p_j')^2}{\tau}, \quad (9a)$$

$$\mathrm{Cov}[p_j', p_k'] = \frac{\sum_{j=0}^{\tau-1} s_j p_j q_j \cdot s_k p_k q_k}{\tau} = \frac{\sum_{j=0}^{\tau-1} p_j' p_k'}{\tau}. \quad (9b)$$

Figure 9 compares Equation 8 to the iterated solution of (2) and (3) as well as to the QLE solution. Equation 8 is accurate only when selection is approximately the same strength as the product of the recombination rates and the period ($\alpha \approx r\tau$ as in Figure 9b), which reflects the order assumptions made in supplementary information (S3). In other cases, the iterated solution of (2) and (3) should be used.

As with the QLE solution (S2.4), Equation 8 indicates that the change in allele frequency at a locus that modifies recombination is proportional to the effect of the modifier on recombination ($\delta r_{j,k|i} p_i q_i$), the inverse of the population size ($2N$), and the square of the selection coefficients at each selected locus (through $\mathrm{Var}[p_j']$ $\mathrm{Var}[p_k']$). Unlike the QLE solution for loose linkage, Equation 8 is not maximized when the allele frequencies remain near $1/2$, because the harmonic mean of the allele frequencies in the denominator is much smaller when the allele frequencies approach zero or one. Consequently, with tighter linkage, selection on a modifier is maximized when there are high-amplitude fluctuations in allele frequencies, such that beneficial alleles reach low frequencies (where drift is stronger) as well as intermediate frequencies (where selection on a modifier is more effective). Selection on the modifier is also strong-

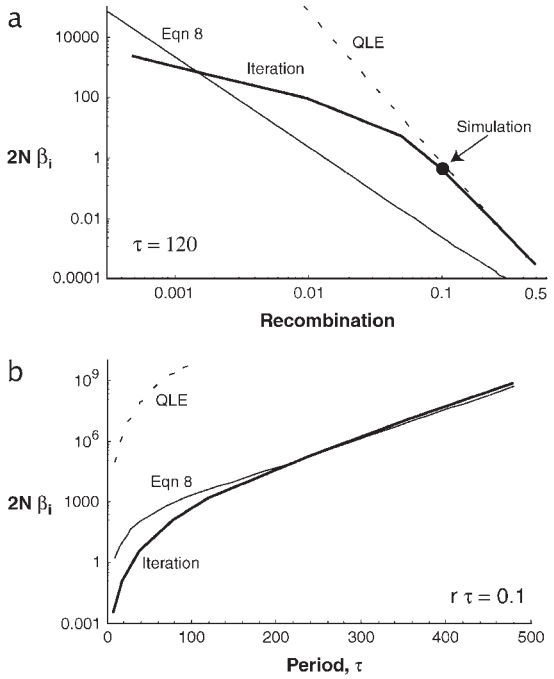

FIGURE 9.—The selection gradient on a modifier of recombination multiplied by the population size, $2N\beta_i$, at steady state under sinusoidal selection, plotted on a log-log scale. The *x*-axis is the recombination rate between adjacent loci, assumed to lie in the order *ijk*, with $r = r_{ij} = r_{jk}$. The thick curve illustrates the predicted change in the modifier from iteration of (2) and (3). The thin solid curve illustrates (8), which is a low-recombination approximation to the recursion equations. The long dashed curve illustrates the QLE solution (S2.4). The dot marks the simulation result reported in Figure 8a. Parameter values were set to $\alpha_j = \alpha_k = 0.1$, $p_{j0} = p_{k0} = 0.5$, with (a) period $\tau = 120$ across a range of recombination rates and (b) a range of periods, τ, with $r\tau = \alpha = 0.1$. The per generation change in the modifier (thick curves) is much higher when the period of fluctuating selection is longer; in this case, alleles pass through both low and high frequency within each period, increasing the fluctuations in disequilibrium that drive selection for the modifier.

est when changes in allele frequency at the two loci occur at similar times, maximizing $\mathrm{Cov}[p_j', p_k']^2$.

Our analysis indicates that there are two qualitatively different regimes with respect to selection on a modifier of recombination, depending on whether recombination rates are high relative to selection [the QLE regime, (S2.4)] or not [the tight linkage regime, (8)]. In the former case, the disequilibrium generated by drift dissipates rapidly, so that disequilibrium can be approximated as the consequence of recent selection (within a time frame proportional to $1/r$), which is well captured by the current additive genetic variance in fitness. When recombination rates are weak relative to selection, however, disequilibrium generated while alleles are at low frequency persists as the alleles rise to higher frequency, and the current additive genetic variance in fitness no longer determines the change in frequency of the modifier. Instead, the force on the modifier is strongest whenever the harmonic mean allele frequency at the two loci, H_{jk}, is small yet the selected alleles spend time at intermediate frequency (with high $\mathrm{Var}[p_j']$, $\mathrm{Var}[p_k']$, and $\mathrm{Cov}[p_j', p_k']$). We also see a qualitatively different relationship between the recombination rates and the change in the modifier frequency in these two regimes. Equation 8 shows that a modifier of recombination changes in frequency by an amount proportional to r^{-3} for tight linkage, while the QLE analysis (see Equations 7) predicts a change that is proportional to r^{-5}. Thus, while both analyses predict that selection for recombination is strongest under tight linkage, the selection gradient is overestimated by the QLE analysis (Figure 9).

DISCUSSION

By breaking down linkage disequilibria that limit the response of a population to selection, sex and recombination can hasten the rate of adaptive evolution of a population. Consequently, genotypes that are best adapted to a novel environment are more likely to be recombinant and more likely to carry modifiers that increase the recombination rate. As these adaptive genotypes sweep through a population, modifier alleles that increase recombination also rise in frequency. Disequilibria do not, however, always limit the response of a population to selection. In fact, the genetic variance for a trait and consequently the response to selection are increased by positive disequilibria among alleles contributing to that trait. If positive disequilibria predominate, recombination is selected against, because the genetic variance and the rate of adaptation are lower among recombinant individuals. In short, a necessary condition for the evolution of increased recombination in a single unstructured population is that negative disequilibria predominate among favored alleles.

Within an unstructured population, negative disequilibrium is expected to arise under two different scenarios. First, negative disequilibrium develops when there is negative epistasis among selected loci, implying that a genotype carrying multiple beneficial alleles has a lower fitness than the product of the fitnesses of the genotypes carrying only one of the beneficial alleles. Second, negative disequilibrium arises by the combined action of random genetic drift and selection, even in the absence of epistasis. Sampling error during reproduction of a finite population causes some genotypes to become more or less common than expected, generating variance in disequilibria. Selection eliminates positive disequilibria more efficiently than it eliminates negative disequilibria, because negative disequilibrium reduces the genetic variance upon which selection acts. Consequently, the average disequilibrium over loci or over replicate populations becomes negative over time.

In this article, we have demonstrated that the random genetic drift caused by sampling in finite populations

selects for increased recombination even in populations that are initially in linkage equilibrium and that are subject to selection acting independently upon all loci. Our method keeps track of the perturbations in genetic associations caused by genetic drift and acted upon by selection, using recursion equations for the first and second moments of these perturbations. The main assumptions of the method are that (a) recombination is altered by only a small amount by a modifier gene, (b) the population is large, and (c) each allele is initially present in several copies. Although we have simplified the analysis by ignoring epistasis, the method can be extended to include epistasis by explicitly tracking the deterministic trajectory for the disequilibrium D_{jk}. Using recursions for the moments describing the distribution of a population around its expected trajectory, the effects of drift and selection can be easily and efficiently studied. Simulations of selection in finite populations require large amounts of computer time, particularly since millions of replicates are needed to obtain accurate measurements of the frequency of a modifier. The simulations that we performed (see figures) indicate that Equations 2 and 3 provide accurate estimates of the average change that occurs at a modifier locus, as long as alleles are always present in multiple copies. It should be kept in mind that we used strong selection ($s = 0.1$) and a strong modifier ($\delta r_{j,k|i} = r_{jk}/2$) in the simulations to magnify the amount of indirect selection acting on the modifier locus. The approximations should be more accurate with weaker selection and weaker modifiers, although it then becomes exceedingly difficult to detect an effect on the modifier with only two selected loci and one modifier locus. Nevertheless, the method provides a promising route to predicting the evolution of recombination when multiple loci of weak effect underlie the selected trait, using the approximation that each pair of selected loci independently asserts an influence on the modifier(s) (see Barton 1995a).

We have calculated the expected change in frequency of the modifier caused by random linkage disequilibria with selected loci, presented in the figures using the selection gradient, $\beta_i = \Delta p_i/(\delta r_{j,k|i} p_i q_i)$. For a modifier of given effect, the effective selection acting on a modifier of recombination is $s_e = \beta_i \delta r_{j,k|i}$. Although absolute values of s_e are small, they can still be much stronger than the effect of random drift, since $2Ns_e$ can be large. For example, for $s = 0.1$, $r_{ij} = r_{jk} = 0.1$, $2N = 10{,}000$, $p_0 = 0.01$, the effective selection is at most $s_e = 0.008\ \delta r_{j,k|i}$ (on the basis of the simulations reported in Figure 6a), but $2Ns_e$ is 80 $\delta r_{j,k|i}$. Even stronger selection for recombination is expected when beneficial alleles start at lower frequency and/or when recombination rates are lower than the selection coefficients (see Figures 7 and 9 and Equation 8). This argument is complicated by the extra stochasticity introduced by the hitchhiking process itself, which reduces the effective population size witnessed by the modifier locus; this effect can be substantial when the rate of selective sweeps is comparable to the recombination rate (Barton 2000, Equation 4).

The QLE approximation for the change in the modifier (7, or more accurately S2.4) shows that when selection is weaker than recombination, the advantage of recombination is proportional to the product of the additive variance in fitness at the directly selected loci and decreases steeply with recombination. This implies that when we average over selective sweeps that are scattered across the genome, the main contribution will come from selection on loci that are most closely linked to the modifier. We cannot use the QLE results to calculate this average, however, because the QLE approximation breaks down for tight linkage. However, it is possible to calculate this average assuming weak selection using integral solutions similar to (A2.4) and (A2.5); this leads to an expression proportional to the square of the total variance in fitness (N. H. Barton, unpublished results).

By comparing Figure 6a and 6b, it can be seen that a modifier allele that increases recombination rates is most strongly selected in smaller populations. Although our approach breaks down when the population size becomes so small that some alleles are present in only a few copies (specifically, approximating Equation 1 to order $1/N$ becomes inadequate), we have shown through simulation that randomly generated disequilibria generate strong selection for increased recombination in populations of small to intermediate size (Otto and Barton 2001). In fact, we found that modifiers of recombination were more influenced by disequilibria generated by drift than by disequilibria generated by epistasis unless the population size was fairly large. Directional selection experiments, such as the one performed by Korol and Iliadi (1994) in *D. melanogaster* (see Otto and Barton 2001 for further review), have demonstrated that recombination rates can increase substantially in small populations subject to strong selection. Our work has shown that the evolution of recombination in such cases can be explained by stochastically generated disequilibria that, over time, accumulate to oppose the action of selection. Furthermore, the stochastic advantage of recombination can even favor the evolution of sex in the face of a twofold cost of sex as long as modifiers increase the amount of sex by small amounts (Otto and Lenormand 2002).

For large populations, our results indicate that selection on recombination should be inversely proportional to the population size. Thus, while drift in the presence of selection can account for increased recombination rates in populations of small size, it would appear to provide a weak basis for explaining the maintenance of sex and recombination in natural populations whose effective sizes are tens of thousands or more. There are three reasons, however, to believe that drift might drive the evolution of sex and recombination even in large populations. First, if population sizes tend to be small

when conditions are harsh and selection is strong, the advantage of breaking down randomly generated disequilibria could drive the evolution of sex and recombination during those rare periods in time that selection is intense. Second, the cumulative selective force on a modifier of recombination might be substantial if there are many loci responding to selection within the genome. Indeed, simulation results reported by Iles *et al.* (2003) demonstrate that selection acting on a modifier of recombination increases as a function of the number of selected loci, even when the initial additive genetic variance in fitness is held constant. Third, and most importantly, the forces described in this article can be substantial even in very large populations as long as those populations are spatially structured (see Martin *et al.* 2005). That is, the disequilibrium generated by drift in the presence of selection depends strongly on the local (deme) size and not just the total size of a population. Because every population is finite, spatial structure is common and selection acts at multiple loci throughout a genome, and because drift in the presence of selection causes an accumulation of good alleles on bad genetic backgrounds, the ubiquity of sex and recombination might well be explained by the fact that genetic mixing allows beneficial alleles in different individuals to be brought together, as initially proposed by Morgan (1913), Fisher (1930), and Muller (1932).

We are grateful to Aneil Agrawal, Mark Kirkpatrick, Thomas Lenormand, Guillaume Martin, Marcy Uyenoyama, and three anonymous reviewers for comments on the manuscript. Financial support for this research was provided by the Biotechnology and Biological Science Research Council, the Natural Environment Research Council, and the Natural Sciences and Engineering Research Council of Canada.

LITERATURE CITED

Altenberg, L., and M. W. Feldman, 1987 Selection, generalized transmission and the evolution of modifier genes. I. The reduction principle. Genetics **117:** 559–572.

Barton, N. H., 1995a A general model for the evolution of recombination. Genet. Res. **65:** 123–144.

Barton, N. H., 1995b Linkage and the limits to natural selection. Genetics **140:** 821–841.

Barton, N. H., 1998 The effect of hitch-hiking on neutral genealogies. Genet. Res. **72:** 123–133.

Barton, N. H., 2000 Genetic hitchhiking. Philos. Trans. R. Soc. Lond. B **355:** 1553–1562.

Barton, N. H., and M. Turelli, 1991 Natural and sexual selection on many loci. Genetics **127:** 229–255.

Charlesworth, B., 1976 Recombination modification in a fluctuating environment. Genetics **83:** 181–195.

Charlesworth, B., 1990 Mutation-selection balance and the evolutionary advantage of sex and recombination. Genet. Res. **55:** 199–221.

Charlesworth, B., 1993 Directional selection and the evolution of sex and recombination. Genet. Res. **61:** 205–224.

de Visser, J. A. G. M., R. F. Hoekstra and H. van den Ende, 1996 The effect of sex and deleterious mutations on fitness in *Chlamydomonas*. Proc. R. Soc. Lond. Ser. B **263:** 193–200.

de Visser, J. A. G. M., R. F. Hoekstra and H. van den Ende, 1997 Test of interaction between genetic markers that affect fitness in *Aspergillus niger*. Evolution **51:** 1499–1505.

Elena, S. F., and R. E. Lenski, 1997 Test of synergistic interactions among deleterious mutations in bacteria. Nature **390:** 395–398.

Ewens, W. J., 1979 *Mathematical Population Genetics*. Springer-Verlag, Berlin.

Feldman, M. W., 1972 Selection for linkage modification. I. Random mating populations. Theor. Popul. Biol. **21:** 430–439.

Feldman, M. W., F. B. Christiansen and L. D. Brooks, 1980 Evolution of recombination in a constant environment. Proc. Natl. Acad. Sci. USA **77:** 4838–4841.

Felsenstein, J., 1974 The evolutionary advantage of recombination. Genetics **78:** 737–756.

Felsenstein, J., 1988 Sex and the evolution of recombination, pp. 74–86 in *The Evolution of Sex*, edited by R. E. Michod and B. R. Levin. Sinauer Press, Sunderland, MA.

Felsenstein, J., and S. Yokoyama, 1976 The evolutionary advantage of recombination. II. Individual selection for recombination. Genetics **83:** 845–859.

Fisher, R. A., 1930 *The Genetical Theory of Natural Selection*. Oxford University Press, Oxford.

Hill, W. G., and A. Robertson, 1966 The effect of linkage on the limits to artificial selection. Genet. Res. **8:** 269–294.

Iles, M. M., K. Walters and C. Cannings, 2003 Recombination can evolve in large finite populations given selection on sufficient loci. Genetics **165:** 333–337.

Keightley, P. D., 1996 Nature of deleterious mutation load in *Drosophila*. Genetics **144:** 1993–1999.

Keightley, P. D., and A. Caballero, 1997 Genomic mutation rates for lifetime reproductive output and lifespan in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA **94:** 3823–3827.

Keightley, P. D., and A. Eyre-Walker, 2000 Deleterious mutations and the evolution of sex. Science **290:** 331–333.

Kimura, M., 1965 Attainment of quasi-linkage equilibrium when gene frequencies are changing by natural selection. Genetics **52:** 875–890.

Kimura, M., and T. Maruyama, 1966 The mutational load with epistatic gene interactions in fitness. Genetics **54:** 1303–1312.

Kondrashov, A. S., 1982 Selection against harmful mutations in large sexual and asexual populations. Genet. Res. **40:** 325–332.

Kondrashov, A. S., 1984 Deleterious mutations as an evolutionary factor. I. The advantage of recombination. Genet. Res. **44:** 199–217.

Kondrashov, A. S., 1988 Deleterious mutations and the evolution of sexual reproduction. Nature **336:** 435–441.

Kondrashov, A. S., 1993 Classification of hypotheses on the advantage of amphimixis. J. Hered. **84:** 372–387.

Korol, A. B., and K. G. Iliadi, 1994 Recombination increase resulting from directional selection for geotaxis in Drosophila. Heredity **72:** 64–68.

Latter, B. D. H., 1965 The response to artificial selection due to autosomal genes of large effect. II. The effects of linkage on limits to selection in finite populations. Aust. J. Biol. Sci. **18:** 1009–1023.

Martin, G., S. P. Otto and T. Lenormand, 2005 Selection for recombination in structured populations. Genetics (in press).

Maynard Smith, J., 1968 Evolution in sexual and asexual populations. Am. Nat. **102:** 469–473.

Maynard Smith, J., 1978 *The Evolution of Sex*. Cambridge University Press, Cambridge, UK.

Maynard Smith, J., 1980 Selection for recombination in a polygenic model. Genet. Res. **35:** 269–277.

Maynard Smith, J., 1988 Selection for recombination in a polygenic model: the mechanism. Genet. Res. **51:** 59–63.

Morgan, T. H., 1913 *Heredity and Sex*. Columbia University Press, New York.

Mukai, T., 1969 The genetic structure of natural populations of *Drosophila melanogaster*. VII. Synergistic interactions of spontaneous mutant polygenes controlling viability. Genetics **61:** 749–761.

Muller, H. J., 1932 Some genetic aspects of sex. Am. Nat. **66:** 118–138.

Nagylaki, T., 1993 The evolution of multilocus systems under weak selection. Genetics **134:** 627–647.

Otto, S. P., and N. H. Barton, 1997 The evolution of recombination: removing the limits to natural selection. Genetics **147:** 879–906.

Otto, S. P., and N. H. Barton, 2001 Selection for recombination in small populations. Evolution **55:** 1921–1931.

Otto, S. P., and M. W. Feldman, 1997 Deleterious mutations, vari-

able epistatic interactions, and the evolution of recombination. Theor. Popul. Biol. **51:** 134–147.

Otto, S. P., and T. Lenormand, 2002 Resolving the paradox of sex and recombination. Nat. Rev. Genet. **2:** 252–261.

Qureshi, A. W., 1963 *A Monte Carlo Evaluation of the Role of Finite Population Size and Linkage in Response to Continuous Mass Selection.* Statistical Laboratory, Iowa State University, Ames, IA.

Rice, W. R., 2002 Experimental tests of the adaptive significance of sexual recombination. Nat. Rev. Genet. **3:** 241–251.

Seager, R. D., and F. J. Ayala, 1982 Chromosome interactions in *Drosophila melanogaster.* I. Viability studies. Genetics **102:** 467–483.

Seager, R. D., F. J. Ayala and R. W. Marks, 1982 Chromosome interactions in *Drosophila melanogaster.* II. Total fitness. Genetics **102:** 485–502.

Stephan, W., T. H. E. Wiehe and M. W. Lenz, 1992 The effect of strongly selected substitutions on neutral polymorphism: analytical results based on diffusion theory. Theor. Popul. Biol. **41:** 237–254.

Szathmary, E., 1993 Do deleterious mutations act synergistically? Metabolic control theory provides a partial answer. Genetics **133:** 127–132.

Turner, J. R. G., and M. H. Williamson, 1968 Population size, natural selection, and the genetic load. Nature **218:** 700.

West, S. A., A. D. Peters and N. H. Barton, 1998 Testing for epistasis between deleterious mutations. Genetics **149:** 435–444.

Wolfram, S., 1991 *Mathematica.* Addison Wesley, New York.

Communicating editor: M. K. Uyenoyama

APPENDIX A: EXACT EQUATIONS FOR TWO AND THREE LOCI

We use the methods of Barton and Turelli (1991) to derive exact recursions for two selected loci (labeled j, k), and a modifier of small effect at locus i. We census immediately after juveniles have been randomly sampled from the propagules produced by the previous generation. After the census, haploid populations undergo selection, random mating, and meiosis to produce the next generation of haploid propagules, where we assume that the number of potential propagules is so large that these processes can be treated deterministically. The model also applies to a diploid population, as long as fitness is the product of the fitness contribution of each haplotype. For diploids, the life cycle involves a census of diploid juveniles, selection, meiosis, syngamy of gametes to produce diploid propagules, and random sampling of N juveniles. Equations A1.2 of Barton (1995a) differ slightly, because diploid fitness was assumed instead to be the sum of the fitness contribution of each haplotype. We assume that the modifier allele P_i increases recombination between loci j and k by a small amount $\delta r_{j,k|i}$. For ease of analysis, we follow Barton (1995a) and set the recombination rate between loci j and k to $r_{jk} + 2q_i\delta r_{j,k|i}$ in P_iP_i individuals, $r_{jk} + (q_i - p_i)\delta r_{j,k|i}$ in P_iQ_j individuals, and $r_{jk} - 2\,p_i\delta r_{j,k|i}$ in Q_jQ_j individuals. (Note that the *effect* of the modifier, measured by the difference between genotypes, is not frequency dependent.) Let r_{ij} and r_{ik} be the recombination rates between the modifier locus and the selected loci, j and k, respectively, and let r_{ijk} be the probability of recombination between any of the loci i, j, k; these terms are used only to determine the offspring of heterozygous P_iQ_j individuals and so their values in other modifier genotypes are immaterial. Disequilibrium terms involving the modifier locus (D_{ij}, D_{ik}, D_{ijk}) will be dominated by the indirect effects of changing recombination between the selected loci and will be proportional to $\delta r_{j,k|i}$. In the following, we assume that the modifier is weak and keep only linear-order terms in $\delta r_{j,k|i}$.

Let $\phi_j = 1 + s_j(p_j - q_j)/2$ and $\phi_k = 1 + s_k(p_k - q_k)/2$ measure the mean fitness of each locus considered separately. Accounting also for the disequilibrium within a population, the mean fitness is equal to

$$\overline{W} = \phi_j\phi_k + s_js_kD_{jk}. \tag{A1}$$

The three-locus recursions are

$$p_j^* = p_j + \frac{s_jp_jq_j\phi_k + s_k(2 - \phi_j)D_{jk}}{\overline{W}} + \zeta_j \tag{A2a}$$

$$D_{jk}^* = \frac{(1 - r_{jk})(1 - s_j^2/4)(1 - s_k^2/4)D_{jk}}{\overline{W}^2} + \zeta_{jk} \tag{A2b}$$

$$p_i^* = p_i + \frac{(s_j\phi_kD_{ij} + s_k\phi_jD_{ik} + s_js_kD_{ijk})}{\overline{W}} + \zeta_i \tag{A2c}$$

$$D_{ij}^* = \frac{(1 - r_{ij})(1 - s_j^2/4)(\phi_k^2D_{ij} - s_k^2D_{ik}D_{jk} + s_k\phi_kD_{ijk})}{\overline{W}^2} + \zeta_{ij} \tag{A2d}$$

$$\begin{aligned} D_{ijk}^* = {} & -p_iq_i\delta r_{j,k|i}D_{jk}\frac{(1 - s_j^2/4)(1 - s_k^2/4)}{\overline{W}^2} \\ & + (D_{ijk}(\phi_j\phi_k - s_js_kD_{jk}) - 2D_{jk}(s_j\phi_kD_{ij} + s_k\phi_jD_{ik}))\frac{(1 - s_j^2/4)(1 - s_k^2/4)(1 - r_{ijk})}{\overline{W}^3} + \zeta_{ijk}, \end{aligned} \tag{A2e}$$

where $r_{ijk} = (r_{ij} + r_{ik} + r_{jk})/2$ for any gene order with and without interference. p_k^* and D_{ik}^* can be obtained from (A2a) and (A2d), respectively, by interchanging subscripts j and k. These equations assume only that $\delta r_{j,k|i}$ is small and are valid for strong selection (s_j, s_k) and strong linkage disequilibrium between the selected loci (D_{jk}). (A2) can

be derived from the three-locus recursion equations presented in Feldman (1972). A supplementary Mathematica package is available that derives (A2) as well as the other main results given in appendixes a and b.

We assume that $2N$ haploid (or N diploid) juveniles are sampled from the propagules produced by the parental generation according to a multinomial distribution as in the standard Wright-Fisher model (Ewens 1979). The moments of the multinomial distribution can therefore be used to determine the expected values of the perturbations in (A2) and the covariances between them. Although there is no expected change in allele frequencies due to drift, $E[\zeta_i] = E[\zeta_j] = E[\zeta_k] = 0$, the expected change in linkage disequilibrium is $E[\zeta_{jk}] = -D_{jk}/(2N)$. Because disequilibrium terms involving the modifier locus are proportional to $\delta r_{j,k|i}$, the expected amount of drift in these terms will be of the order $\delta r_{j,k|i}/N$, which is assumed to be very small and is ignored. Thus, we need to quantify the effects of drift only on the following variances and covariances, which are written to order $(1/N)$:

$$\mathrm{var}(\zeta_j) = \frac{p_j q_j}{2N}, \quad \mathrm{cov}(\zeta_j, \zeta_k) = \frac{D_{jk}}{2N}, \quad \mathrm{cov}(\zeta_j, \zeta_{jk}) = \frac{-D_{jk}(p_j - q_j)}{2N},$$

$$\mathrm{var}(\zeta_k) = \frac{p_k q_k}{2N}, \quad \mathrm{cov}(\zeta_k, \zeta_{jk}) = \frac{-D_{jk}(p_k - q_k)}{2N},$$

$$\mathrm{var}(\zeta_{jk}) = \frac{(p_j q_j p_k q_k + (p_j - q_j)(p_k - q_k)D_{jk} - D_{jk}^2)}{2N}.$$

Because $E[\zeta_a]$ is either 0 or order $1/N$, the $E[\zeta_a\zeta_b]$ required in Equation 3 are, to order $1/N$, equal to the $\mathrm{cov}(\zeta_a, \zeta_b)$ given here. Note that the allele frequencies and disequilibrium in these terms are evaluated after the parental generation produces propagules. When selection is weak, however, the allele frequencies and disequilibrium from the previous census can be used to obtain leading-order approximations.

These derivations apply for any amount of disequilibrium, allowing the method to be extended to models where there is a deterministic source of disequilibrium (*e.g.*, epistasis or an initially high level of disequilibrium). In the text and appendix b, we assume that the disequilibria remain $O(1/N)$ throughout the process. In this case, $E[\zeta_{jk}]$ and the covariance terms become negligible. In the numerical iterations of (2) and (3), however, no assumptions are made about the order of the disequilibrium, so that the Mathematica package (available upon request) can be used regardless of the initial level of disequilibrium.

APPENDIX B: WEAK SELECTION APPROXIMATIONS FOR TWO SELECTED LOCI

When selection is weak (s_j and s_k of order s, where $s \ll 1$), the recursions for the perturbations given by (1) can be further approximated by

$$\delta p_j^* = \delta p_j(1 - s_j(p_j - q_j)) + s_k\delta D_{jk} - s_j\delta p_j^2 + \zeta_j + O(s^2) \tag{B1a}$$

$$\delta D_{jk}^* = \delta D_{jk}(1 - r_{jk})(1 - s_j(p_j - q_j) - s_k(p_k - q_k)) - 2(1 - r_{jk})(s_j\delta p_j\delta D_{jk} + s_k\delta p_k\delta D_{jk} + s_j s_k\delta D_{jk}^2) + \zeta_{jk} + O(s^2). \tag{B1b}$$

Throughout this appendix, values for allele P_k can be obtained from values given for allele P_j by interchanging subscripts j and k. In (B1b), we have included the leading-order terms for each of the perturbations with respect to s even when these are $O(s^2)$. These lower-order terms are critical to the initial development of disequilibria when δD_{jk} is zero.

Taking expectations as in (4) and (5), assuming that the perturbations including the disequilibria are $O(N^{-1})$, ignoring terms that are $O(N^{-2})$, and using the results of appendix a for the contributions due to drift gives

$$E[\delta p_j^*] = (1 - s_j(p_j - q_j))E[\delta p_j] + s_k E[\delta D_{jk}] - s_j E[\delta p_j^2] \tag{B2a}$$

$$E[\delta p_j^{*2}] = (1 - 2s_j(p_j - q_j))E[\delta p_j^2] + 2s_k E[\delta p_j\delta D_{jk}] + \frac{p_j q_j}{2N} \tag{B2b}$$

$$E[\delta D_{jk}^{*2}] = (1 - r_{jk})^2(1 - 2s_j(p_j - q_j) - 2s_k(p_k - q_k))E[\delta D_{jk}^2] + \frac{p_j q_j p_k q_k}{2N} \tag{B2c}$$

$$E[\delta p_j^*\delta D_{jk}^*] = (1 - r_{jk})(1 - 2s_j(p_j - q_j) - s_k(p_k - q_k))E[\delta p_j\delta D_{jk}] + s_k(1 - r_{jk})E[\delta D_{jk}^2] \tag{B2d}$$

$$E[\delta D_{jk}^*] = (1 - r_{jk})(1 - s_j(p_j - q_j) - s_k(p_k - q_k))E[\delta D_{jk}] - 2(1 - r_{jk})(s_j E[\delta p_j\delta D_{jk}] + s_k E[\delta p_k\delta D_{jk}] + s_j s_k E[\delta D_{jk}^2]). \tag{B2e}$$

Supplementary information (S2) describes similar weak selection recursions for the three-locus case.

Equations B2 can be solved explicitly. First, consider loose linkage ($r_{jk} \gg s_j, s_k$). In this case, the disequilibrium rapidly approaches a QLE distribution, with $E[\delta D_{jk}^2]$ and $E[\delta D_{jk}]$ changing slowly relative to the action of recombination. One can therefore set $E[\delta D_{jk}^*]$ to $E[\delta D_{jk}]$, etc., in (B2) and solve for the QLE values. For the disequilibrium:

$$E[\delta D_{jk}^2] \approx \frac{p_j q_j p_k q_k}{2N r_{jk}(2 - r_{jk})} \tag{B3a}$$

$$E[\delta p_j \delta D_{jk}] \approx \frac{s_k(1 - r_{jk})}{r_{jk}} E[\delta D_{jk}^2] \approx \frac{s_k(1 - r_{jk}) p_j q_j p_k q_k}{2N r_{jk}^2(2 - r_{jk})} \tag{B3b}$$

$$\begin{aligned} E[\delta D_{jk}] &\approx -\frac{2(1 - r_{jk})}{r_{jk}}(s_j E[\delta p_j \delta D_{jk}] + s_k E[\delta p_k \delta D_{jk}] + s_j s_k E[\delta D_{jk}^2]) \\ &\approx -\frac{2(1 - r_{jk})}{r_{jk}} s_j s_k E[\delta D_{jk}^2]\left(\frac{2(1 - r_{jk})}{r_{jk}} + 1\right) \\ &\approx -\frac{2 s_j s_k p_j q_j p_k q_k (1 - r_{jk})}{2N r_{jk}^3}. \end{aligned} \tag{B3c}$$

(B3c) shows that negative disequilibrium is generated on average in a finite population subject to multiplicative selection. The magnitude of this disequilibrium is quite small with loose linkage but increases rapidly as r_{jk} becomes small. It also increases with the strength of selection acting on loci j and k and with the amount of genetic variability at these loci. Note that the covariance between allele frequencies and linkage disequilibrium (B3b) always contributes more than the variance in linkage disequilibrium to the accumulation of negative linkage disequilibrium by a factor $2(1 - r_{jk})/r_{jk}$.

With tight linkage ($1/N \ll s \approx r_{jk} \ll 1$), disequilibrium decays slowly over time, and we can no longer assume quasi-linkage equilibrium. We can, however, move from the discrete recursion equations to continuous-time approximations, which can be more readily solved. We introduce the method in general and then apply the solution to (B2).

In the ODE approach, the discrete recursion equations (B2) all have the form

$$x[t + 1] = (1 + C)(1 + F[t])x[t] + G[t], \tag{B4a}$$

where C is a constant involving recombination rates, $F[t]$ is a function of time involving terms such as $-s(p - q)$, and $G[t]$ is a function of time involving the expectations, $E[\ ]$. When selection and recombination are rare, the change in $x[t]$ over a short time period can be approximated by $x[t + \Delta t] - x[t]$, where $x[t + \Delta t]$ is obtained from (B4a) under the assumption that the amount of selection and recombination in a shorter interval of time, Δt, scales with Δt as $s\Delta t$ and $r\Delta t$. Taking the limit of $(x[t + \Delta t] - x[t])/\Delta t$ as Δt goes to zero yields an ordinary differential equation:

$$\frac{dx}{dt} = (c + f[t])x[t] + g[t], \tag{B4b}$$

where c, f, and g are used to denote C, F, and G to linear order in s and r, using a Taylor series that ignores any cross products or higher-order terms. The solution to differential equations of the form (B4b) is

$$x[t] = e^{ct} e^{\int_{t_1=0}^{t} f[t_1]dt_1}\left(x[0] + \int_{t_2=0}^{t} \frac{g[t_2]}{e^{ct_2} e^{\int_{t_3=0}^{t_2} f[t_3]dt_3}} dt_2\right). \tag{B4c}$$

The integrals involving $f[t]$ in (B4c) can be solved by noting that, when selection is weak and frequency independent, selection can be defined according to the change in allele frequency that it causes: $p' = dp/dt = spq$. When $f[t] = -s(p - q)$, for example, applying this definition and integrating yields

$$\int_{t_1=0}^{t} -s(p - q)\, dt_1 = \int_{t_1=0}^{t} \frac{q'}{q} + \frac{p'}{p} dt_1 = \ln\left(\frac{p_t q_t}{p_0 q_0}\right).$$

Letting $v[t]$ be the product of the appropriate allele frequencies (depending on which terms are included in $f[t]$), (B4c) becomes

$$x[t] = v[t]\left(e^{ct}\frac{x[0]}{v[0]} + \int_{t_2=0}^{t} \frac{e^{c(t - t_2)} g[t_2]}{v[t_2]} dt_2\right). \tag{B4d}$$

The solution (B4d) applies regardless of how selection varies over time. Barton (2000) used a similar method to find the net hitchhiking effect of fluctuating selection on linked neutral loci and gave an approximation valid for tight linkage ($r \ll s$; his Equation 4).

Following this procedure and assuming that terms of order $r_{jk}s_j$ are negligible relative to r_{jk} and s_j, we obtain the following weak selection approximations:

$$E[\delta D_{jk}^2] = (p_j q_j p_k q_k)^2 \left(\frac{e^{-2r_{jk}t} E_0[\delta D_{jk}^2]}{(p_{j0} q_{j0} p_{k0} q_{k0})^2} + \int_{t_1=0}^{t} \frac{e^{-2r_{jk}(t-t_1)}\, dt_1}{2N \cdot p_j q_j p_k q_k} \right) \tag{B5a}$$

$$E[\delta p_j \delta D_{jk}] = (p_j q_j)^2 p_k q_k \left(\frac{e^{-r_{jk}t} E_0[\delta p_j \delta D_{jk}]}{(p_{j0} q_{j0})^2 p_{k0} q_{k0}} + \int_{t_1=0}^{t} \frac{s_k E[\delta D_{jk}^2]\, e^{-r_{jk}(t-t_1)}\, dt_1}{(p_j q_j)^2 p_k q_k} \right) \tag{B5b}$$

$$E[\delta D_{jk}] = p_j q_j p_k q_k \left(\frac{e^{-r_{jk}t} E_0[\delta D_{jk}]}{p_{j0} q_{j0} p_{k0} q_{k0}} - 2 \int_{t_1=0}^{t} \frac{(s_j E[\delta p_j \delta D_{jk}] + s_k E[\delta p_k \delta D_{jk}])\, e^{-r_{jk}(t-t_1)}\, dt_1}{p_j q_j p_k q_k} \right) \tag{B5c}$$

We have focused on the effects of drift on the disequilibrium; analogous equations can be obtained for the effects on allele frequencies. Note that if allele frequencies increase from some low value faster than linkage disequilibrium is dissipated by recombination (s_j, $s_k > r_{jk}$), the initial distribution ($E_0[\ ]$) will make a substantial contribution. Even when disequilibrium is initially absent, substantial disequilibrium can build up if the beneficial alleles start at low frequency, because periods with low values of $p_j q_j p_k q_k$ contribute disproportionately to the integrals in (B5).

Evolution of recombination due to random drift

NICHOLAS H. BARTON & SARAH P. OTTO

Genetics (2005)

Supplementary Information 1: Differentials around the deterministic trajectory

The effects of random genetic drift on the dynamics described in Appendix 1 depend on the first and second partial derivatives of the recursions in (A1.2) with respect to small perturbations in the allele frequencies and disequilibria. For example, if the allele frequency departs from the expected trajectory by a small amount in one generation (δp_j), it will contribute to the perturbation in the next generation (δp_j^*) by an amount approximately equal to $\frac{\partial \delta p_j^*}{\partial \delta p_j}\,\delta p_j$, as described in (1). The partial derivative can be obtained by replacing p_j with $p_j + \delta p_j$ and p_j^* with $p_j^* + \delta p_j^*$ in (A1.2a), and taking the derivative with respect to δp_j. Because these differentials are multiplied in (2) and (3) by perturbations that are O(1/*N*), we will drop terms that are functions of the perturbations (δp_j, δD_{jk}, etc.) within the differentials, thereby assuming that the perturbations including the linkage disequilibrium remain small throughout the process. In other words, we evaluate the partial derivatives along the deterministic trajectory expected in an infinite population. This approximation is made to simplify the presentation and is not made in the numerical analyses reported in the figures.

For the two-locus model, the first differentials of δp_j^* evaluated along the deterministic trajectory are:

$$\frac{\partial \delta p_j^*}{\partial \delta p_j} = \frac{(1 - s_j^2/4)}{\phi_j^2} \qquad \frac{\partial \delta p_j^*}{\partial \delta p_k} = 0 \qquad \frac{\partial \delta p_j^*}{\partial \delta D_{jk}} = \frac{s_k(1 - s_j^2/4)}{\phi_j^2 \phi_k}$$

The second differentials of δp_j^* are:

$$\frac{\partial^2 \delta p_j^*}{\partial \delta p_j^2} = -\frac{2s_j(1 - s_j^2/4)}{\phi_j^3} \qquad \frac{\partial^2 \delta p_j^*}{\partial \delta p_j \, \partial \delta p_k} = 0 \qquad \frac{\partial^2 \delta p_j^*}{\partial \delta p_j \, \partial \delta D_{jk}} = -\frac{2s_j s_k(1 - s_j^2/4)}{\phi_j^3 \phi_k}$$

$$\frac{\partial^2 \delta p_j^*}{\partial \delta p_k^2} = 0 \qquad \frac{\partial^2 \delta p_j^*}{\partial \delta p_k \, \partial \delta D_{jk}} = -\frac{s_k^2(1 - s_j^2/4)}{\phi_j^2 \phi_k^2}$$

$$\frac{\partial^2 \delta p_j^*}{\partial \delta D_{jk}^2} = -\frac{2s_j s_k^2(1 - s_j^2/4)}{\phi_j^3 \phi_k^2}$$

The first and second differentials of δp_k^* may be obtained by interchanging *j* and *k* throughout the above. The first differentials of δD_{jk}^* are:

$$\frac{\partial \delta D_{jk}^*}{\partial \delta p_j} = 0 \qquad \frac{\partial \delta D_{jk}^*}{\partial \delta p_k} = 0 \qquad \frac{\partial \delta D_{jk}^*}{\partial \delta D_{jk}} = \frac{(1 - r_{jk})(1 - s_j^2/4)(1 - s_k^2/4)}{\phi_j^2 \phi_k^2}$$

The second differentials of δD_{jk}^* are given by:

$$\frac{\partial^2 \delta D_{jk}^*}{\partial \delta p_j^2} = 0 \qquad \frac{\partial^2 \delta D_{jk}^*}{\partial \delta p_j \, \partial \delta p_k} = 0 \qquad \frac{\partial^2 \delta D_{jk}^*}{\partial \delta p_j \, \partial \delta D_{jk}} = -\frac{2s_j(1 - r_{jk})(1 - s_j^2/4)(1 - s_k^2/4)}{\phi_j^3 \phi_k^2}$$

$$\frac{\partial^2 \delta D^*_{jk}}{\partial \delta p_k^2} = 0 \qquad \frac{\partial^2 \delta D^*_{jk}}{\partial \delta p_k \, \partial \delta D_{jk}} = -\frac{2 s_k (1 - r_{jk})(1 - s_j^2/4)(1 - s_k^2/4)}{\phi_j^2 \phi_k^3}$$

$$\frac{\partial^2 \delta D^*_{jk}}{\partial \delta D_{jk}^2} = -\frac{4 s_j s_k (1 - r_{jk})(1 - s_j^2/4)(1 - s_k^2/4)}{\phi_j^3 \phi_k^3}$$

With the modifier locus, the three-locus model requires additional first and second derivatives such as $\frac{\partial \delta p_i^*}{\partial \delta p_i}$. These are calculated as above and are available in the supplementary *Mathematica* package.

Supplementary Information 2: Weak selection approximation with a modifier

In this appendix, we describe the weak selection approximation for the recursions involving a modifier and then determine the quasi-linkage equilibrium (QLE) value for the change in the modifier under the assumption that recombination rates are higher than the selection coefficients. The exact recursions describing the expected perturbations and their covariances for the three-locus system were derived by substituting the differentials described in (S1) into (2) and (3), dropping terms of $O(N^{-2})$. Assuming that selection is weak, the recursions for the expectations become:

$$\begin{aligned}\mathrm{E}[\,\delta p_i^*\,] \approx\ & \mathrm{E}[\delta p_i] + s_j\,\mathrm{E}[\delta D_{ij}] + s_k\,\mathrm{E}[\delta D_{ik}] + s_j s_k\,\mathrm{E}[\delta D_{ijk}] \\ & - \big(s_j^2\,\mathrm{E}[\delta p_j\,\delta D_{ij}] + s_j^2 s_k\,\mathrm{E}[\delta p_j\,\delta D_{ijk}] + s_k^2\,\mathrm{E}[\delta p_k\,\delta D_{ik}] + s_j s_k^2\,\mathrm{E}[\delta p_k\,\delta D_{ijk}] \\ & + s_j^2 s_k\,\mathrm{E}[\delta D_{jk}\,\delta D_{ij}] + s_j s_k^2\,\mathrm{E}[\delta D_{jk}\,\delta D_{ik}] + s_j^2 s_k^2\,\mathrm{E}[\delta D_{jk}\,\delta D_{ijk}]\big), \end{aligned} \tag{S2.1a}$$

$$\begin{aligned}\mathrm{E}[\,\delta C_{ij}^*\,] \approx\ & (1-r_{ij})\,\big(1-s_j(p_j-q_j)\big)\,\mathrm{E}[\delta D_{ij}] + s_k\,(1-r_{ij})\,\mathrm{E}[\delta D_{ijk}] \\ & - (1-r_{ij})\,\big(2s_j\,\mathrm{E}[\delta p_j\,\delta D_{ij}] + 2\,s_j s_k\,\mathrm{E}[\delta p_j\,\delta D_{ijk}] + s_k^2\,\mathrm{E}[\delta p_k\,\delta D_{ijk}] \\ & + 2\,s_j s_k\,\mathrm{E}[\delta D_{jk}\,\delta D_{ij}] + s_k^2\,\mathrm{E}[\delta D_{jk}\,\delta D_{ik}] + 2\,s_j s_k^2\,\mathrm{E}[\delta D_{jk}\,\delta D_{ijk}]\big) \end{aligned} \tag{S2.1b}$$

$$\begin{aligned}\mathrm{E}[\,\delta D_{ijk}^*\,] \approx\ & -\delta r_{j,k|i}\,p_i q_i\,\big(1-s_j(p_j-q_j)-s_k(p_k-q_k)\big)\mathrm{E}[\delta D_{jk}] + (1-r_{ijk})\big(1-s_j(p_j-q_j)-s_k(p_k-q_k)\big)\mathrm{E}[\delta D_{ijk}] \\ & + 2\delta r_{j,k|i}\,p_i q_i\,\big(s_j\,\mathrm{E}[\delta p_j\,\delta D_{jk}] + s_k\,\mathrm{E}[\delta p_k\,\delta D_{jk}] + s_j s_k\,\mathrm{E}[\delta\,D_{jk}^2]\big) \end{aligned}$$

$$- 2(1-r_{ijk}) \left(s_j \, \mathrm{E}[\delta p_j \, \delta D_{ijk}] + s_k \, \mathrm{E}[\delta p_k \, \delta D_{ijk}] + s_j \, \mathrm{E}[\delta D_{jk} \, \delta D_{ij}] \right.$$

$$\left. + s_k \, \mathrm{E}[\delta D_{jk} \, \delta D_{ik}] + 2 \, s_j \, s_k \, \mathrm{E}[\delta D_{jk} \, \delta D_{ijk}] \right) \qquad \text{(S2.1c)}$$

Throughout this appendix, terms involving locus k can be obtained from terms involving locus j by interchanging subscripts j and k, e.g. $\mathrm{E}[\delta D^*_{ik}]$ may thus be obtained from (S2.1b). Here we have kept leading-order terms in the selection coefficients for each expectation, even if they appear to be lower order, because such terms play a role in the initial build up of the disequilibria. We continue to ignore terms that are order $\delta r^2_{j,k|i}$, such as $\mathrm{E}[\delta p_i \, \delta D_{ij}]$ and $\delta r_{j,k|i} \, \mathrm{E}[\delta p_i \, \delta D_{jk}]$, thereby assuming that the modifier's effect on recombination is weak.

To evaluate the recursions (S2.1), the following covariances are needed, which constitute a closed system of equations:

$$\mathrm{E}[\delta D^*_{jk} \, \delta D^*_{ijk}] \approx (1-r_{jk}) \left(1-2s_j(p_j-q_j)-2s_k(p_k-q_k)\right) \left((1-r_{ijk}) \, \mathrm{E}[\delta D_{jk} \, \delta D_{ijk}] - \delta r_{j,k|i} \, p_i q_i \, \mathrm{E}[\delta D^2_{jk}]\right) \qquad \text{(S2.2a)}$$

$$\mathrm{E}[\delta D^*_{jk} \, \delta D^*_{ij}] \approx (1-r_{jk})(1-r_{ij}) \left(\left(1-2s_j(p_j-q_j)-s_k(p_k-q_k)\right) \mathrm{E}[\delta D_{jk} \, \delta D_{ij}] + s_k \, \mathrm{E}[\delta D_{jk} \, \delta D_{ijk}]\right) \qquad \text{(S2.2b)}$$

$$\mathrm{E}[\delta p^*_j \, \delta D^*_{ijk}] \approx \left(1-2s_j(p_j-q_j)-s_k(p_k-q_k)\right) \left((1-r_{ijk}) \, \mathrm{E}[\delta p_j \, \delta D_{ijk}] - \delta r_{j,k|i} \, p_i q_i \, \mathrm{E}[\delta p_j \, \delta D_{jk}]\right)$$

$$+ s_k \, (1-r_{ijk}) \, \mathrm{E}[\delta D_{jk} \, \delta D_{ijk}] - s_k \, \delta r_{j,k|i} \, p_i q_i \, \mathrm{E}[\delta D^2_{jk}] \qquad \text{(S2.2c)}$$

$$\mathrm{E}[\,\delta p^*_j\,\delta D^*_{ij}\,] \approx (1-r_{ij})\,\big(1-2s_j(p_j-q_j)\big)\,\mathrm{E}[\delta p_j\,\delta D_{ij}]$$

$$+\,(1-r_{ij})\,\big(s_k\,\mathrm{E}[\delta p_j\,\delta D_{ijk}] + s_k\,\mathrm{E}[\delta D_{jk}\,\delta D_{ij}] + s_k^2\,\mathrm{E}[\delta D_{jk}\,\delta D_{ijk}]\big) \qquad \text{(S2.2d)}$$

$$\mathrm{E}[\,\delta p^*_j\,\delta D^*_{ik}\,] \approx (1-r_{ik})\,\big(1-s_j(p_j-q_j)-s_k(p_k-q_k)\big)\,\mathrm{E}[\delta p_j\,\delta D_{ik}]$$

$$+\,(1-r_{ik})\,\big(s_j\,\mathrm{E}[\delta p_j\,\delta D_{ijk}] + s_k\,\mathrm{E}[\delta D_{jk}\,\delta D_{ik}] + s_j\,s_k\,\mathrm{E}[\delta D_{jk}\,\delta D_{ijk}]\big) \qquad \text{(S2.2e)}$$

These equations demonstrate that the covariances involving the modifier locus are ultimately driven by variance in linkage disequilibrium ($\mathrm{E}[\delta D^2_{jk}]$), either directly as in (S2.2a, c) or indirectly through the covariance terms $\mathrm{E}[\delta p_j\,\delta D_{jk}]$ and $\mathrm{E}[\delta p_k\,\delta D_{jk}]$.

Under weak selection, recursions (S2.1) and (S2.2) can be iterated to determine the effects of drift and selection on the modifier. When selection is weak relative to the recombination rates, (S2.1b-c) and (S2.2) rapidly approach quasi-linkage equilibrium (QLE) values, which can be determined by setting, for example, $\mathrm{E}[\,\delta D^*_{jk}\,\delta D^*_{ijk}\,]$ to $\mathrm{E}[\delta D_{jk}\,\delta D_{ijk}]$ and solving. The results depend on recombination terms such as $1-(1-r_a)(1-r_b)$, which we write as $R_{a,b}$ to save space. Using the two-locus QLE results (A2.3), the covariance terms (S2.2) at QLE are:

$$\mathrm{E}[\delta D_{jk}\,\delta D_{ijk}] = -\frac{\delta r_{j,k\backslash i}\; p_i q_i (1-r_{jk})\, E[\delta D^2_{jk}]}{R_{ijk,\,jk}}$$

$$= -\frac{\delta r_{j,k\backslash i}\,(1-r_{jk})\; p_i q_i p_j q_j p_k q_k}{2N\, r_{jk}\,(2-r_{jk})\, R_{ijk,\,jk}} \qquad \text{(S2.3a)}$$

$$\mathrm{E}[\delta D_{jk}\,\delta D_{ij}] \;=\; \frac{s_k(1-r_{jk})(1-r_{ij})\;E[\delta D_{jk}\,\delta D_{ijk}]}{R_{ij,\,jk}}$$

$$=\; -\frac{\delta r_{j,k\backslash i}\;s_k\,(1-r_{jk})^2\,(1-r_{ij})\,p_iq_ip_jq_jp_kq_k}{2N\,r_{jk}\,(2-r_{jk})\,R_{ij,\,jk}R_{ijk,\,jk}} \qquad \text{(S2.3b)}$$

$$\mathrm{E}[\delta p_j\,\delta D_{ijk}] \;=\; \frac{s_k\,(1-r_{ijk})\,E[\delta D_{jk}\,\delta D_{ijk}] - \delta r_{j,k\backslash i}\;p_iq_i\,E[\delta p_j\,\delta D_{jk}] - \delta r_{j,k\backslash i}\;s_k\;p_iq_i\,E[\delta D_{jk}^2]}{r_{ijk}}$$

$$=\; -\frac{\delta r_{j,k\backslash i}\;s_k\left(1-(1-r_{ijk})(1-r_{jk})^2\right)p_iq_ip_jq_jp_kq_k}{2N\,r_{jk}^2\,r_{ijk}\,(2-r_{jk})\,R_{ijk,\,jk}} \qquad \text{(S2.3c)}$$

$$\mathrm{E}[\delta p_j\,\delta D_{ij}] \;=\; \frac{(1-r_{ij})\left(s_kE[\delta p_j\,\delta D_{ijk}] + s_kE[\delta D_{jk}\,\delta D_{ij}] + s_k^2E[\delta D_{jk}\,\delta D_{ijk}]\right)}{r_{ij}} \qquad \text{(S2.3d)}$$

$$=\; -\frac{\delta r_{j,k\backslash i}\,s_k^2\;p_iq_ip_jq_jp_kq_k\,(1-r_{ij})\left(r_{jk}(2-r_{jk})\left(1-(1-r_{ij})(1-r_{jk})(1-r_{ijk})\right) + r_{ij}r_{ijk}(1-r_{jk})\right)}{2N\,r_{ij}r_{ijk}r_{jk}^2(2-r_{jk})R_{ij,\,jk}R_{ijk,\,jk}}$$

$$\mathrm{E}[\delta p_j\,\delta D_{ik}] \;=\frac{(1-r_{ik})\left(s_jE[\delta p_j\,\delta D_{ijk}] + s_kE[\delta D_{jk}\,\delta D_{ik}] + s_js_kE[\delta D_{jk}\,\delta D_{ijk}]\right)}{r_{ik}} \qquad \text{(S2.3e)}$$

$$-\frac{\delta r_{j,k\backslash i}\,s_js_k\;p_iq_ip_jq_jp_kq_k\,(1-r_{ik})\left(r_{jk}(2-r_{jk})\left(1-(1-r_{ik})(1-r_{jk})(1-r_{ijk})\right) + r_{ik}r_{ijk}(1-r_{jk})\right)}{2N\,r_{ik}r_{ijk}r_{jk}^2(2-r_{jk})R_{ik,\,jk}R_{ijk,\,jk}}$$

The QLE value of $\mathrm{E}[\delta D_{ijk}]$ can be found by substituting (S2.3) into (S2.1c) and solving for $\mathrm{E}[\delta D_{ijk}]$. This value can then be used along with (S2.3) to find $\mathrm{E}[\delta D_{ij}]$ (and, by interchanging *j*

and k, $E[\delta D_{ik}]$) from (S2.1b). Finally, all of these QLE values may be used in (S2.1a) to solve for the change in modifier frequency over time, yielding:

$$E[\delta p_i^*] = E[\delta p_i] + \frac{\delta r_{j,k|i}\ p_i q_i V_j V_k\ \ f(r)}{8N\ r_{jk}^5} \tag{S2.4}$$

where $V_j = 2\,s_j^2\,p_j q_j$ and $V_k = 2\,s_k^2\,p_k q_k$ and where $f(r)$ is a positive, if complicated, function of the recombination rates:

$$f(r) = \frac{r_{jk}^2}{R_{ij,\,jk} R_{ik,\,jk} R_{ijk,\,jk}\ r_{ij}^2 r_{ik}^2 r_{ijk}^2\,(2 - r_{jk})} \Big\{ 4\ R_{ij,\,jk} R_{ik,\,jk} R_{ijk,\,jk} \left(R_{ij,ik} - r_{ijk}\right) r_{ij} r_{ik} r_{jk}\ \left(2 - r_{jk}\right)$$
$$+ 2\ r_{ijk} r_{jk}^2\ R_{ij,ik} \Big(R_{ij,\,jk} R_{ik,\,jk} \left(r_{ij} + r_{ik}\right)\left(2 - r_{jk}\right) + r_{ij} r_{ik} \left(1 - r_{jk}\right)\left(R_{ij,ik} - r_{ijk}\right) + \left(2 - r_{jk}\right) X \Big)$$
$$+ r_{ijk}^2 \left(2 - r_{jk}\right) \Big(2 R_{ij,ik}^2 R_{ij,\,jk} R_{ik,\,jk} \left(1 - r_{jk}\right) r_{jk} + R_{ij,ik} r_{jk}^2 X + \left(1 - r_{jk}\right) r_{ij} r_{ik} \Big(2 R_{ij,ik} \Big(r_{ij} r_{ik} \left(1 - r_{jk}^2\right) + r_{jk}^2 \Big) + R_{ij,ik}^2 r_{jk} + 2X \Big) \Big) \Big\}$$

with $X = \left(1 - r_{jk}\right) r_{jk}\ \left(1 - r_{ij}\right) r_{ij}\ \left(1 - r_{ik}\right) r_{ik}$.

To better understand equation (S2.4), we assume that the recombination rates are small, all of the same order, but larger than the selection coefficients as required for QLE (s_j, $s_k << r_{jk}$, r_{ij}, r_{ik}, $r_{ijk} << 1$) and that there is no genetic interference among loci. Without loss of generality, we also assume that locus k is the right-most locus and define $\alpha = r_{ik}/r_{jk}$. Equation (S2.4) may then be simplified by taking a Taylor Series approximation. Keeping only the leading-order terms with respect to selection and recombination, we arrive at (7). This approximation allows us to

replace the complicated function $f(r)$ in (S2.4) with the function $g(\alpha)$ given by (7b) for gene order *jik* and (7c) for gene order *ijk*.

Supplementary Information 3: Steady state solution for the change in a modifier under fluctuating selection

In this appendix, we consider fluctuating selection with period, τ. Such selection could be sinusoidal, it could involve bouts of selection favoring one allele followed by bouts favoring the other allele (selection following a square wave), or it could involve periods of no selection interspersed with periods of positive and negative selection. We assume that selection is weak, that each cycle is identical, and that alleles return to the same frequency after each cycle. As in Appendix 2, we could replace the recursion equations involving a modifier (S2.1 – S2.2) with differential equations, whose solutions have the form of equation (A2.4b). We will take a slightly different tack, however, and analyze the discrete recursion equations directly.

As mentioned earlier, each of the variables has a recursion of the form (A2.4a). Furthermore, the term (1+f[t]) term involves quantities such as $1 - s_t(p_t - q_t)$. Under the assumption of weak selection, $p_{t+1} \approx p_t + s_t p_t q_t$, which can be rearranged to give $1 - s_t(p_t - q_t) \approx \frac{p_{t+1}q_{t+1}}{p_t q_t}$. Thus, we can replace (1+f[t]) in (A2.4a) with the term $v[t+1]/v[t]$, where again $v[t]$ is the product of the appropriate allele frequencies at time t. For example, in the recursion equation for E[δC_{jk}^2], (A2.2c), we approximate $(1 - 2\, s_j(p_j - q_j) - 2\, s_k(p_k - q_k))$ at time t with $v[t+1]/v[t]$ where $v[t] = (p_j q_j p_k q_k)^2$. The solution to (A2.4a) is then:

$$x[t] = v[t]\left((1+c)^t \frac{x[0]}{v[0]} + \sum_{j=0}^{t-1} \frac{(1+c)^{t-1-j}\, g[j]}{v[j+1]}\right). \qquad \text{(S3.1a)}$$

Using (S3.1a), we can evaluate $x[t]$ after one full cycle $x[t+\tau]$. If the system is at steady state, the two values should coincide, $x[t] = x[t+\tau]$. We can use this fact to solve (S3.1a) for its steady state value at point, t, within a cycle:

$$x[t] = \frac{v[t]\sum_{j=0}^{\tau-1}\frac{(1+c)^{\tau-j}\,g[j]}{v[j+1]}}{1-(1+c)^{\tau}}. \qquad \text{(S3.1b)}$$

To our knowledge, there is no explicit solution to the sum in (S3.1b) or to its continuous-time analog. To proceed, we assume that the recombination rates and, hence, c are small (so that $c\,\tau << 1$). Furthermore, we assume that selection is sufficiently weak that $v[t+1]$ can be approximated by $v[t]$. These assumptions allow us to rewrite (S3.1b) as:

$$x[t] = \frac{v[t]\sum_{j=0}^{\tau-1}\frac{g[j]}{v[j]}}{c\,\tau}, \qquad \text{(S3.1c)}$$

which is simply:

$$x[t] = \frac{v[t]\,\text{Mean}\left[\frac{g[j]}{v[j]}\right]}{c}. \qquad \text{(S3.1d)}$$

where $\text{Mean}\left[\frac{g[j]}{v[j]}\right]$ is evaluated over the course of one cycle.

Application of (S3.1d) to the recursions (A2.2), (S2.1), and (S2.2) is straightforward because $\mathrm{Mean}\left[\frac{g[j]}{v[j]}\right]$ is often a quantity that can be readily determined from the dynamics of selection. For example, to estimate $E_t[\delta D_{jk}^2]$, we require the mean of $\frac{1}{2N\, p_j q_j p_k q_k}$, which is given by $2N$ times the harmonic mean, H_{jk}, of $p_j q_j p_k q_k$:

$$H_{jk} = \frac{\tau}{\sum_{j=0}^{\tau-1} \frac{1}{p_j q_j p_k q_k}} . \tag{S3.2}$$

Plugging this into (S3.1d) gives the approximate solution:

$$E_t[\delta C_{jk}^2] = \frac{\left(p_j q_j p_k q_k\right)^2}{2N\, H_{jk}\left(2r_{jk}\right)} \tag{S3.3a}$$

In other cases, $\mathrm{Mean}\left[\frac{g[j]}{v[j]}\right]$ is approximately zero. For example, to estimate $E_t[\delta p_j\, \delta D_{jk}]$, we require the mean of $\frac{s_k E_t[\delta D_{jk}^2]}{(p_j q_j)^2 p_k q_k}$. Plugging in (S3.3a) leaves us with $s_k p_k q_k$ times the constant $1/\left(2N\, H_{jk}\, 2r_{jk}\right)$. Under our assumption that selection is weak, $s_k p_k q_k$ describes the change in allele frequency per generation. For the allele frequencies to return to their original state, the average value of $s_k p_k q_k$ must be approximately zero.

As a final example, because $E_t[\delta p_j \, \delta D_{jk}]$ and, by symmetry, $E_t[\delta p_k \, \delta D_{jk}]$ are zero, $E_t[\delta D_{jk}]$ can be approximated from the mean value of $-2\frac{s_j s_k E_t[\delta D_{jk}^2]}{p_j q_j p_k q_k} = -2\frac{s_j s_k p_j q_j p_k q_k}{2N\, H_{jk}\left(2r_{jk}\right)}$.

Defining the covariance in allele frequency changes at the two loci using (9b) then gives us an approximation for the disequilibrium:

$$E_t[\delta D_{jk}] = -\frac{p_j q_j p_k q_k \; Cov[p'_j, p'_k]}{2N\, H_{jk} r_{jk}^2} \qquad \text{(S3.3b)}$$

Repeated application of this method to all of the recursions yields the approximation (8) for the change in the modifier allele (see supplementary *Mathematica* package).

Aside: If one were to apply this procedure to find the steady state solution to the differential equations of the form A2.4b, one would also get $E_t[\delta p_j \, \delta D_{jk}] = 0$. This poses a problem, however, when solving for the disequilibrium, $E_t[\delta D_{jk}]$. Without $E_t[\delta p_j \, \delta D_{jk}]$ and $E_t[\delta p_k \, \delta D_{jk}]$, there is no driving term in (A2.5c) to generate disequilibrium in the continuous-time model. As a result, the predicted change in the modifier is zero. This problem does not arise in the discrete-time model, because $E_t[\delta D_{jk}^2]$ also contributes to the development of disequilibrium in (A2.2e). Thus, the weak selection and weak recombination approximations made in the continuous-time version of the model fail to capture the phenomenon of interest. In general, we have found the behaviour of the continuous-time model with weak fluctuating selection and weak recombination to be extremely sensitive to the assumptions made.